```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         July  3, 2001, 10:33:07 ; Search time 34.09 Seconds
                                          (without alignments)
                                          550.660 Million cell updates/sec

Title:          US-09-508-418-2
Perfect score:  2844
Sequence:       1 MTTTPIANHPNIFTHQSSSS..........EGAYEVASEVTGFLSRYAYK 548

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES

                %
Result          Query
  No.   Score   Match Length DB  ID                   Description
-------------------------------------------------------------------------
    1    2821    99.2    548  1  PPOC_TOBAC           O24163 nicotiana t
    2    2075    73.0    537  1  PPOC_ARATH           P55826 arabidopsis
    3     541    19.0    471  1  PPOX_MYXXA           P56601 myxococcus
    4     505    17.8    504  1  PPOM_TOBAC           O24164 nicotiana t
    5     485    17.1    470  1  PPOX_BACSU           P32397 bacillus su
    6   375.5    13.2    527  1  PPOX_PROFR           O32434 propionibac
    7   363.5    12.8    477  1  PPOX_HUMAN           P50336 homo sapien
    8   362.5    12.7    477  1  PPOX_MOUSE           P51175 mus musculu
    9   302.5    10.6    490  1  YAM7_SCHPO           Q10062 schizosacch
   10     221     7.8    426  1  PPOX_MYCTU           O53230 mycobacteri
   11   216.5     7.6    451  1  PPOX_MYCLE           Q50008 mycobacteri
   12     189     6.6    539  1  PPOX_YEAST           P40012 saccharomyc
   13     141     5.0    448  1  AOFH_MYCTU           O53320 mycobacteri
   14     139     4.9    570  1  CRTI_NARPS           Q40406 narcissus p
   15     134     4.7    522  1  AOF_ONCMY            P49253 oncorhynchu
   16   132.5     4.7    527  1  AOFA_BOVIN           P21398 bos taurus
   17     131     4.6    566  1  CRTI_ARATH           Q07356 arabidopsis
   18   125.5     4.4    417  1  Y4AB_RHISN           P55349 rhizobium s
   19   125.5     4.4    472  1  CRTI_SYNY3           P29273 synechocyst
   20   125.5     4.4    489  1  CBP1_CANAL           P31225 candida alb
   21   123.5     4.3    526  1  AOFA_RAT             P21396 rattus norv
   22     123     4.3    582  1  CRTI_CAPAN           P80093 capsicum an
   23     122     4.3    471  1  Y782_SYNY3           Q55629 synechocyst
   24   119.5     4.2    478  1  PUO_MICRU            P40974 micrococcus
   25     118     4.1    571  1  CRTI_MAIZE           P49086 zea mays (m
   26   117.5     4.1    469  1  P49_STRLI            P06108 streptomyce
   27     116     4.1    583  1  CRTI_LYCES           P28554 lycopersico
   28   114.5     4.0   1106  1  ITA7_RAT             Q63258 rattus norv
   29     114     4.0    531  1  ACHC_ACHFU           P35903 achatina fu
   30   112.5     4.0    570  1  CRTI_SOYBN           P28553 glycine max
   31     111     3.9    508  1  CRTI_STRSE           P54971 streptomyce
   32   109.5     3.9    527  1  AOFA_HUMAN           P21397 homo sapien
   33   109.5     3.9    908  1  SRCA_RABIT           P13666 oryctolagus
```

```
   34     109     3.8   1181  1  ITA7_HUMAN           Q13683 homo sapien
   35   108.5     3.8   1062  1  CERU_MOUSE           Q61147 mus musculu
   36     108     3.8    504  1  A37C_DROME           P18487 drosophila
   37   107.5     3.8    516  1  ER13_ARATH           O65403 arabidopsis
   38     107     3.8    405  1  SAOX_CORS1           P40875 corynebacte
   39     107     3.8   1179  1  ITA7_MOUSE           Q61738 mus musculu
   40   106.5     3.7    630  1  FIG1_MOUSE           O09046 mus musculu
   41   105.5     3.7    872  1  IP3L_RAT             P42335 rattus norv
   42     105     3.7    755  1  TR2M_AGRT3           P25017 agrobacteri
   43   104.5     3.7    466  1  AMID_PSEPU           O69768 pseudomonas
   44   104.5     3.7    517  1  ER12_ARATH           O65402 arabidopsis
   45   104.5     3.7    524  1  CRTI_RHOCA           P17054 rhodobacter

                              ALIGNMENTS

RESULT  1
PPOC_TOBAC
ID   PPOC_TOBAC     STANDARD;      PRT;   548 AA.
AC   O24163;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO I)
DE   (PROTOPORPHYRINOGEN IX OXIDASE ISOZYME I) (PPX I).
GN   PPXI.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. SR1;
RX   MEDLINE=97385200; PubMed=9238074;
RA   Lermontova I., Kruse E., Mock H.-P., Grimm B.;
RT   "Cloning and characterization of a plastidal and a mitochondrial
RT   isoform of tobacco protoporphyrinogen IX oxidase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:8895-8900(1997).
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
CC   -!- COFACTOR: FAD (BY SIMILARITY).
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS, AND
CC       CHLOROPHYLL SYNTHESIS.
CC   -!- SUBUNIT: HOMODIMER; CONTAINS ONE FAD PER HOMODIMER (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED IN EXPANDING PREMATURE LEAVES.
CC       DECREASED EXPRESSION IN OLDEST LEAVES. EXPRESSED AT VERY LOW LEVEL
CC       IN ROOTS.
CC   -!- INDUCTION: OSCILLATING EXPRESSION DURING DIURNAL GROWTH. MAXIMAL
CC       EXPRESSION IN THE DARK PERIOD.
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y13465; CAA73865.1; -.
KW   Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW   Transit peptide; Flavoprotein; FAD; Chloroplast;
KW   Chlorophyll biosynthesis.
FT   TRANSIT       1     50       CHLOROPLAST (POTENTIAL).
FT   CHAIN        51    548       PROTOPORPHYRINOGEN OXIDASE.
FT   NP_BIND      78     83       FAD (POTENTIAL).
SQ   SEQUENCE   548 AA;  59230 MW;  66892E78FB8A3E30 CRC64;
```

```
  Query Match             99.2%;  Score 2821;  DB 1;  Length 548;
  Best Local Similarity   99.3%;  Pred. No. 1.4e-205;
  Matches  544;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60

Qy       61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120

Qy      121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMS 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db      121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLAFFDLMS 180

Qy      181 IPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMK 240
            |||||||||| ||||||||||||||||||||||||||||||||||||||| |||||||||
Db      181 IPGKLRAGFGAIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYAGDPSKLSMK 240

Qy      241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300

Qy      301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360

Qy      361 SVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420
            |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db      361 SVAAADALSNFYYPPVGAVTITYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420

Qy      421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480

Qy      481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540

Qy      541 FLSRYAYK 548
            ||||||||
Db      541 FLSRYAYK 548


RESULT   2
PPOC_ARATH
ID   PPOC_ARATH     STANDARD;      PRT;   537 AA.
AC   P55826;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO).
GN   PPOX OR AT4G01690 OR T15B16.13.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97136707; PubMed=8982084;
RA   Narita S., Tanaka R., Ito T., Okada K., Taketani S., Inokuchi H.;
RT   "Molecular cloning and characterization of a cDNA that encodes
RT   protoporphyrinogen oxidase of Arabidopsis thaliana.";
RL   Gene 182:169-175(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
     STRAIN=CV. COLUMBIA;
     MEDLINE=20083488; PubMed=10617198;
       r K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sehkon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaugnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
CC   -!- COFACTOR: FAD (BY SIMILARITY).
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST (POTENTIAL).
CC   -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE LEAVES AND AT
CC       LOW LEVELS IN THE ROOTS AND FLORAL BUDS.
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D83139; BAA11820.1; -.
DR   EMBL; AL161492; CAB77739.1; -.
KW   Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW   Flavoprotein; FAD; Chloroplast; Transit peptide.
FT   TRANSIT       1      ?       CHLOROPLAST (POTENTIAL).
FT   CHAIN         ?    537       PROTOPORPHYRINOGEN OXIDASE.
FT   NP_BIND      63     68       FAD (POTENTIAL).
SQ   SEQUENCE   537 AA;  57695 MW;  C125B0EE1D3187D5 CRC64;
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         July  3, 2001, 10:00:01 ; Search time 28.95 Seconds
                                          (without alignments)
                                          381.322 Million cell updates/sec

Title:          US-09-508-418-2
Perfect score:  2844
Sequence:       1 MTTTPIANHRNIFTHQSSSS..........EGAYEVASEVTGFLSRYAYK 548

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       193259 seqs, 20144636 residues

Total number of hits satisfying chosen parameters:      193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
              1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
              2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
              3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
              4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
              5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
              6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES

                  %
Result          Query
  No.   Score   Match Length DB  ID                         Description
--------------------------------------------------------------------------
    1    2113    74.3    539  2  US-08-808-931-16           Sequence 16, Appl
    2    2113    74.3    539  3  US-08-808-323-16           Sequence 16, Appl
    3    2113    74.3    539  3  US-09-050-603A-16          Sequence 16, Appl
    4    2113    74.3    539  3  US-09-102-420B-16          Sequence 16, Appl
    5    2075    73.0    537  1  US-08-472-028A-2           Sequence 2, Appli
    6    2075    73.0    537  2  US-08-808-931-2            Sequence 2, Appli
    7    2075    73.0    537  3  US-08-808-323-2            Sequence 2, Appli
    8    2075    73.0    537  3  US-09-050-603A-2           Sequence 2, Appli
    9    2075    73.0    537  3  US-09-102-420B-2           Sequence 2, Appli
   10    2075    73.0    537  4  US-09-071-296-2            Sequence 2, Appli
   11    2055    72.3    543  2  US-08-808-931-12           Sequence 12, Appl
   12    2055    72.3    543  3  US-08-808-323-12           Sequence 12, Appl
   13    2055    72.3    543  3  US-09-050-603A-12          Sequence 12, Appl
   14    2055    72.3    543  3  US-09-102-420B-12          Sequence 12, Appl
   15  2006.5    70.6    536  2  US-08-808-931-20           Sequence 20, Appl
   16  2006.5    70.6    536  3  US-08-808-323-20           Sequence 20, Appl
   17  2006.5    70.6    536  3  US-09-050-603A-20          Sequence 20, Appl
   18  2006.5    70.6    536  3  US-09-102-420B-20          Sequence 20, Appl
   19    1987    69.9    528  2  US-08-808-931-10           Sequence 10, Appl
   20    1987    69.9    528  3  US-08-808-323-10           Sequence 10, Appl
   21    1987    69.9    528  3  US-09-050-603A-10          Sequence 10, Appl
   22    1987    69.9    528  3  US-09-102-420B-10          Sequence 10, Appl
   23    1978    69.5    481  1  US-08-472-028A-6           Sequence 6, Appli
   24    1978    69.5    481  2  US-08-808-931-6            Sequence 6, Appli
   25    1978    69.5    481  3  US-08-808-323-6            Sequence 6, Appli
   26    1978    69.5    481  3  US-09-050-603A-6           Sequence 6, Appli
   27    1978    69.5    481  3  US-09-102-420B-6           Sequence 6, Appli
   28    1978    69.5    483  4  US-09-071-296-6            Sequence 6, Appli
   29  1933.5    68.0    560  2  US-08-808-931-18           Sequence 18, Appl
   30  1933.5    68.0    560  3  US-08-808-323-18           Sequence 18, Appl
   31  1933.5    68.0    560  3  US-09-050-603A-18          Sequence 18, Appl
   32  1933.5    68.0    560  3  US-09-102-420B-18          Sequenc
   33    1823    64.1    440  2  US-08-808-931-24           Sequenc
   34    1823    64.1    440  3  US-08-808-323-24           Sequenc
   35    1823    64.1    440  3  US-09-050-603A-24          Sequenc
   36    1823    64.1    440  3  US-09-102-420B-24          Sequenc
   37    1261    44.3    312  2  US-08-808-931-22           Sequence 22, Appl
   38    1261    44.3    312  3  US-08-808-323-22           Sequence 22, Appl
   39    1261    44.3    312  3  US-09-050-603A-22          Sequence 22, Appl
   40    1261    44.3    312  3  US-09-102-420B-22          Sequence 22, Appl
   41   490.5    17.2    508  1  US-08-472-028A-4           Sequence 4, Appli
   42   490.5    17.2    508  2  US-08-808-931-4            Sequence 4, Appli
   43   490.5    17.2    508  3  US-08-808-323-4            Sequence 4, Appli
   44   490.5    17.2    508  3  US-09-050-603A-4           Sequence 4, Appli
   45   490.5    17.2    508  3  US-09-102-420B-4           Sequence 4, Appli

                              ALIGNMENTS

RESULT   1
US-08-808-931-16
; Sequence 16, Application US/08808931
; Patent No. 5939602
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5939602artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,931
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  539 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-808-931-16

  Query Match             74.3%;  Score 2113;  DB 2;  Length 539;
  Best Local Similarity   79.2%;  Pred. No. 9.4e-202;
  Matches  401;  Conservative   46;  Mismatches   55;  Indels    4;  Gaps    1;

Qy     47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
          :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db     34 FKLRCSLAEGPTISSSKIDGGESSIADCVIVGGGISGLCIAQALATKHRDVASNVIVTEA 93

Qy    103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
          ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db     94 RDRVGGNITTVERDGYLWEEGPNSFQPSDPILTMAVDSGLKDDLVLGDPNAPRFVLWEGK 153

Qy    163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
          ||||||| ||||||||||| |||||||| ||:|| |||:|||||:||||||| ||||| |
Db    154 LRPVPSKPTDLPFFDLMSIAGKLRAGFGAIGIRPPPPGYEESVEEFVRRNLGAEVFERFI 213

Qy    223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
          |||||||| |||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db    214 EPFCSGVYAGDPSKLSMKAAFGRVWKLEEIGGSIIGGTFKTIQERNKTPKPPRDPRLPKP 273

Qy    283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
          |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
Db    274 KGQTVGSFRKGLTMLPEAIANSLGSNVKLSWKLSSITKLGNGGYNLTFETPEGMVSLQSR 333

Qy    343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
          |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db    334 SVVMTIPSHVASNLLHPLSAAAADALSQFYYPPVASVTVSYPKEAIRKECLIDGELKGFG 393

Qy    403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
          |||||:||:||||||||||||||||| |||||||||||| |  |||||| :||| |||||
Db    394 QLHPRSQGIETLGTIYSSSLFPNRAPSGRVLLLNYIGGATNTGILSKTEGELVEAVDRDL 453

Qy    463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
          ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db    454 RKMLINPNAKDPLVLGVRVWPKAIPQFLVGHLDLLDSAKMALRDSGFHGLFLGGNYVSGV 513

Qy    523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
          |||||||||||||:||  |||:||||
Db    514 ALGRCVEGAYEVAAEVKEFLSQYAYK 539

RESULT   2
US-08-808-323-16
; Sequence 16, Application US/08808323
; Patent No. 6018105
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Ward, Eric
;    TITLE OF INVENTION:  Promoters from Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase Genes
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6018105artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      [illegible]PERATING SYSTEM:  PC-DOS/MS-DOS
;      [illegible]TWARE:  PatentIn Release #1.0, Version #1.30
;    [illegible] APPLICATION DATA:
;      [illegible]TION NUMBER:  US/08/808,323
;      [illegible]TE:
```

14

```
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1846
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  539 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-808-323-16

  Query Match             74.3%;  Score 2113;  DB 3;  Length 539;
  Best Local Similarity   79.2%;  Pred. No. 9.4e-202;
  Matches  401;  Conservative   46;  Mismatches   55;  Indels    4;  Gaps    1;

Qy     47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
          :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db     34 FKLRCSLAEGPTISSSKIDGGESSIADCVIVGGGISGLCIAQALATKHRDVASNVIVTEA 93

Qy    103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
          ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db     94 RDRVGGNITTVERDGYLWEEGPNSFQPSDPILTMAVDSGLKDDLVLGDPNAPRFVLWEGK 153

Qy    163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
          ||||||| ||||||||||| |||||||| ||:|| |||:|||||:||||||| ||||| |
Db    154 LRPVPSKPTDLPFFDLMSIAGKLRAGFGAIGIRPPPPGYEESVEEFVRRNLGAEVFERFI 213

Qy    223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
          |||||||| |||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db    214 EPFCSGVYAGDPSKLSMKAAFGRVWKLEEIGGSIIGGTFKTIQERNKTPKPPRDPRLPKP 273

Qy    283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
          |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
Db    274 KGQTVGSFRKGLTMLPEAIANSLGSNVKLSWKLSSITKLGNGGYNLTFETPEGMVSLQSR 333

Qy    343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
          |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db    334 SVVMTIPSHVASNLLHPLSAAAADALSQFYYPPVASVTVSYPKEAIRKECLIDGELKGFG 393

Qy    403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
          |||||:||:||||||||||||||||| |||||||||||| |  |||||| :||| |||||
Db    394 QLHPRSQGIETLGTIYSSSLFPNRAPSGRVLLLNYIGGATNTGILSKTEGELVEAVDRDL 453

Qy    463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
          ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db    454 RKMLINPNAKDPLVLGVRVWPKAIPQFLVGHLDLLDSAKMALRDSGFHGLFLGGNYVSGV 513

Qy    523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
          |||||||||||||:||  |||:||||
Db    514 ALGRCVEGAYEVAAEVKEFLSQYAYK 539

RESULT   3
US-09-050-603A-16
; Sequence 16, Application US/09050603A
; Patent No. 6023012
;  GENERAL INFORMATION:
```

```
                       GenCore version 4.5
             Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          July  3, 2001, 10:00:01 ; Search time 28.95 Seconds
                                              (without alignments)
                                              381.322 Million cell updates/sec

Title:           US-09-508-418-2
Perfect score:   2844
Sequence:        1 MTTTPIANHPNIFTHQSSSS..........EGAYEVASEVTGFLSRYAYK 548

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters:      193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result            Query
  No.     Score   Match  Length DB  ID                       Description
--------------------------------------------------------------------------
   1       2113    74.3    539  2  US-08-808-931-16          Sequence 16, Appl
   2       2113    74.3    539  3  US-08-808-323-16          Sequence 16, Appl
   3       2113    74.3    539  3  US-09-050-603A-16         Sequence 16, Appl
   4       2113    74.3    539  3  US-09-102-420B-16         Sequence 16, Appl
   5       2075    73.0    537  1  US-08-472-028A-2          Sequence 2, Appli
   6       2075    73.0    537  2  US-08-808-931-2           Sequence 2, Appli
   7       2075    73.0    537  3  US-08-808-323-2           Sequence 2, Appli
   8       2075    73.0    537  3  US-09-050-603A-2          Sequence 2, Appli
   9       2075    73.0    537  3  US-09-102-420B-2          Sequence 2, Appli
  10       2075    73.0    537  4  US-09-071-296-2           Sequence 2, Appli
  11       2055    72.3    543  2  US-08-808-931-12          Sequence 12, Appl
  12       2055    72.3    543  3  US-08-808-323-12          Sequence 12, Appl
  13       2055    72.3    543  3  US-09-050-603A-12         Sequence 12, Appl
  14       2055    72.3    543  3  US-09-102-420B-12         Sequence 12, Appl
  15     2006.5    70.6    536  2  US-08-808-931-20          Sequence 20, Appl
  16     2006.5    70.6    536  3  US-08-808-323-20          Sequence 20, Appl
  17     2006.5    70.6    536  3  US-09-050-603A-20         Sequence 20, Appl
  18     2006.5    70.6    536  3  US-09-102-420B-20         Sequence 20, Appl
  19       1987    69.9    528  2  US-08-808-931-10          Sequence 10, Appl
  20       1987    69.9    528  3  US-08-808-323-10          Sequence 10, Appl
  21       1987    69.9    528  3  US-09-050-603A-10         Sequence 10, Appl
  22       1987    69.9    528  3  US-09-102-420B-10         Sequence 10, Appl
  23       1978    69.5    481  1  US-08-472-028A-6          Sequence 6, Appli
  24       1978    69.5    481  2  US-08-808-931-6           Sequence 6, Appli
  25       1978    69.5    481  3  US-08-808-323-6           Sequence 6, Appli
  26       1978    69.5    481  3  US-09-050-603A-6          Sequence 6, Appli
  27       1978    69.5    481  3  US-09-102-420B-6          Sequence 6, Appli
  28       1978    69.5    483  4  US-09-071-296-6           Sequence 6, Appli
  29     1933.5    68.0    560  2  US-08-808-931-18          Sequence 18, Appl
  30     1933.5    68.0    560  3  US-08-808-323-18          Sequence 18, Appl
  31     1933.5    68.0    560  3  US-09-050-603A-18         Sequence 18, Appl
  32     1933.5    68.0    560  3  US-09-102-420B-18         Sequence 18, Appl
  33       1823    64.1    440  2  US-08-808-931-24          Sequence 24, Appl
  34       1823    64.1    440  3  US-08-808-323-24          Sequence 24, Appl
  35       1823    64.1    440  3  US-09-050-603A-24         Sequence 24, Appl
  36       1823    64.1    440  3  US-09-102-420B-24         Sequence 24, Appl
  37       1261    44.3    312  2  US-08-808-931-22          Sequence 22, Appl
  38       1261    44.3    312  3  US-08-808-323-22          Sequence 22, Appl
  39       1261    44.3    312  3  US-09-050-603A-22         Sequence 22, Appl
  40       1261    44.3    312  3  US-09-102-420B-22         Sequence 22, Appl
  41      490.5    17.2    508  1  US-08-472-028A-4          Sequence 4, Appli
  42      490.5    17.2    508  2  US-08-808-931-4           Sequence 4, Appli
  43      490.5    17.2    508  3  US-08-808-323-4           Sequence 4, Appli
  44      490.5    17.2    508  3  US-09-050-603A-4          Sequence 4, Appli
  45      490.5    17.2    508  3  US-09-102-420B-4          Sequence 4, Appli

                              ALIGNMENTS

RESULT   1
US-08-808-931-16
; Sequence 16, Application US/08808931
; Patent No. 5939602
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5939602artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,931
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  539 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-808-931-16

  Query Match             74.3%;  Score 2113;  DB 2;  Length 539;
  Best Local Similarity   79.2%;  Pred. No. 9.4e-202;
  Matches  401;  Conservative   46;  Mismatches   55;  Indels    4;  Gaps    1;

Qy      47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
           :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db      34 FKLRCSLAEGPTISSSKIDGGESSIADCVIVGGGISGLCIAQALATKHRDVASNVIVTEA 93

Qy     103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
           ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db      94 RDRVGGNITTVERDGYLWEEGPNSFQPSDPILTMAVDSGLKDDLVLGDPNAPRFVLWEGK 153

Qy     163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
           ||||||| ||||||||||| |||||||| ||:|| |||:|||||:||||||| ||||| |
Db     154 LRPVPSKPTDLPFFDLMSIAGKLRAGFGAIGIRPPPPGYEESVEEFVRRNLGAEVFERFI 213

Qy     223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
           |||||||| |||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db     214 EPFCSGVYAGDPSKLSMKAAFGRVWKLEEIGGSIIGGTFKTIQERNKTPKPPRDPRLPKP 273

Qy     283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
           |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
Db     274 KGQTVGSFRKGLTMLPEAIANSLGSNVKLSWKLSSITKLGNGGYNLTFETPEGMVSLQSR 333

Qy     343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
           |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db     334 SVVMTIPSHVASNLLHPLSAAAADALSQFYYPPVASVTVSYPKEAIRKECLIDGELKGFG 393

Qy     403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
           |||||:||:||||||||||||||||| |||||||||||| |  |||||| :||| |||||
Db     394 QLHPRSQGIETLGTIYSSSLFPNRAPSGRVLLLNYIGGATNTGILSKTEGELVEAVDRDL 453

Qy     463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
           ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db     454 RKMLINPNAKDPLVLGVRVWPKAIPQFLVGHLDLLDSAKMALRDSGFHGLFLGGNYVSGV 513

Qy     523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
           |||||||||||||:||  |||:||||
Db     514 ALGRCVEGAYEVAAEVKEFLSQYAYK 539

RESULT   2
US-08-808-323-16
; Sequence 16, Application US/08808323
; Patent No. 6018105
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Ward, Eric
;    TITLE OF INVENTION:  Promoters from Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase Genes
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6018105artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,323
;      FILING DATE:
```

```
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1846
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  539 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-808-323-16

  Query Match             74.3%;  Score 2113;  DB 3;  Length 539;
  Best Local Similarity   79.2%;  Pred. No. 9.4e-202;
  Matches  401;  Conservative   46;  Mismatches   55;  Indels    4;  Gaps    1;

Qy      47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
           :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db      34 FKLRCSLAEGPTISSSKIDGGESSIADCVIVGGGISGLCIAQALATKHRDVASNVIVTEA 93

Qy     103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
           ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db      94 RDRVGGNITTVERDGYLWEEGPNSFQPSDPILTMAVDSGLKDDLVLGDPNAPRFVLWEGK 153

Qy     163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
           ||||||| ||||||||||| |||||||| ||:|| |||:|||||:||||||| ||||| |
Db     154 LRPVPSKPTDLPFFDLMSIAGKLRAGFGAIGIRPPPPGYEESVEEFVRRNLGAEVFERFI 213

Qy     223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
           |||||||| |||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db     214 EPFCSGVYAGDPSKLSMKAAFGRVWKLEEIGGSIIGGTFKTIQERNKTPKPPRDPRLPKP 273

Qy     283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
           |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
Db     274 KGQTVGSFRKGLTMLPEAIANSLGSNVKLSWKLSSITKLGNGGYNLTFETPEGMVSLQSR 333

Qy     343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
           |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db     334 SVVMTIPSHVASNLLHPLSAAAADALSQFYYPPVASVTVSYPKEAIRKECLIDGELKGFG 393

Qy     403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
           |||||:||:||||||||||||||||| |||||||||||| |  |||||| :||| |||||
Db     394 QLHPRSQGIETLGTIYSSSLFPNRAPSGRVLLLNYIGGATNTGILSKTEGELVEAVDRDL 453

Qy     463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
           ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db     454 RKMLINPNAKDPLVLGVRVWPKAIPQFLVGHLDLLDSAKMALRDSGFHGLFLGGNYVSGV 513

Qy     523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
           |||||||||||||:||  |||:||||
Db     514 ALGRCVEGAYEVAAEVKEFLSQYAYK 539

RESULT   3
US-09-050-603A-16
; Sequence 16, Application US/09050603A
; Patent No. 6023012
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6023012artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/050,603A
;      FILING DATE:  30-MAR-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  539 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-09-050-603A-16

  Query Match             74.3%;  Score 2113;  DB 3;  Length 539;
  Best Local Similarity   79.2%;  Pred. No. 9.4e-202;
  Matches  401;  Conservative   46;  Mismatches   55;  Indels    4;  Gaps    1;

Qy     47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
          :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db     34 FKLRCSLAEGPTISSSKIDGGESSIADCVIVGGGISGLCIAQALATKHRDVASNVIVTEA 93

Qy    103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
          ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db     94 RDRVGGNITTVERDGYLWEEGPNSFQPSDPILTMAVDSGLKDDLVLGDPNAPRFVLWEGK 153

Qy    163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
          ||||||| ||||||||||| |||||||| ||:|| |||:|||||:||||||| ||||| |
Db    154 LRPVPSKPTDLPFFDLMSIAGKLRAGFGAIGIRPPPPGYEESVEEFVRRNLGAEVFERFI 213

Qy    223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
          |||||||| |||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db    214 EPFCSGVYAGDPSKLSMKAAFGRVWKLEEIGGSIIGGTFKTIQERNKTPKPPRDPRLPKP 273
```

```
Qy    283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
          |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
Db    274 KGQTVGSFRKGLTMLPEAIANSLGSNVKLSWKLSSITKLGNGGYNLTFETPEGMVSLQSR 333

Qy    343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
          |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db    334 SVVMTIPSHVASNLLHPLSAAAADALSQFYYPPVASVTVSYPKEAIRKECLIDGELKGFG 393

Qy    403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
          |||||:||:||||||||||||||||| |||||||||||| |  |||||| :||| |||||
Db    394 QLHPRSQGIETLGTIYSSSLFPNRAPSGRVLLLNYIGGATNTGILSKTEGELVEAVDRDL 453

Qy    463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
          ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db    454 RKMLINPNAKDPLVLGVRVWPKAIPQFLVGHLDLLDSAKMALRDSGFHGLFLGGNYVSGV 513

Qy    523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
          |||||||||||||:||  |||:||||
Db    514 ALGRCVEGAYEVAAEVKEFLSQYAYK 539

RESULT    4
US-09-102-420B-16
; Sequence 16, Application US/09102420B
; Patent No. 6084155
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
;    TITLE OF INVENTION:  OXIDASE ("PROTOX")
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6084155artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/102,420B
;      FILING DATE:  22-JUN-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/059,164
;      FILING DATE:  13-APR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/050,603
;      FILING DATE:  30-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/126,430
;      FILING DATE:  11-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/472,028
;      FILING DATE:  06-JUN-1995
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847/CIP3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  539 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-09-102-420B-16

  Query Match             74.3%;  Score 2113;  DB 3;  Length 539;
  Best Local Similarity   79.2%;  Pred. No. 9.4e-202;
  Matches  401;  Conservative   46;  Mismatches   55;  Indels    4;  Gaps    1;

Qy     47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
          :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db     34 FKLRCSLAEGPTISSSKIDGGESSIADCVIVGGGISGLCIAQALATKHRDVASNVIVTEA 93

Qy    103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
          ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db     94 RDRVGGNITTVERDGYLWEEGPNSFQPSDPILTMAVDSGLKDDLVLGDPNAPRFVLWEGK 153

Qy    163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
          ||||||| ||||||||||| |||||||| ||:|| ||| |||||:||||||| ||||| |
Db    154 LRPVPSKPTDLPFFDLMSIAGKLRAGFGAIGIRPPPPGYEESVEEFVRRNLGAEVFERFI 213

Qy    223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
          |||||||| |||||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db    214 EPFCSGVYAGDPSKLSMKAAFGRVWKLEEIGGSIIGGTFKTIQERNKTPKPPRDPRLPKP 273

Qy    283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
          |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
Db    274 KGQTVGSFRKGLTMLPEAIANSLGSNVKLSWKLSSITKLGNGGYNLTFETPEGMVSLQSR 333

Qy    343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
          |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db    334 SVVMTIPSHVASNLLHPLSAAAADALSQFYYPPVASVTVSYPKEAIRKECLIDGELKGFG 393

Qy    403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
          |||||:||:||||||||||||||||| |||||||||||| |  |||||| :||| |||||
Db    394 QLHPRSQGIETLGTIYSSSLFPNRAPSGRVLLLNYIGGATNTGILSKTEGELVEAVDRDL 453

Qy    463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
          ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db    454 RKMLINPNAKDPLVLGVRVWPKAIPQFLVGHLDLLDSAKMALRDSGFHGLFLGGNYVSGV 513

Qy    523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
          |||||||||||||:||  |||:||||
Db    514 ALGRCVEGAYEVAAEVKEFLSQYAYK 539

RESULT   5
US-08-472-028A-2
; Sequence 2, Application US/08472028A
; Patent No. 5767373
;  GENERAL INFORMATION:
;    APPLICANT:  Ward, Eric R
;    APPLICANT:  Volrath, Sandra
;    TITLE OF INVENTION:  Manipulation of Protoporphyrinogen
;    TITLE OF INVENTION:  Oxidase Enzyme Activity in Eukaryotic Organisms
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ciba-Geigy Corporation
;      STREET:  7 Skyline Drive
;      CITY:  Hawthorne
;      STATE:  NY
```

```
;      COUNTRY:  USA
;      ZIP:  10532
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/472,028A
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Elmer, James Scott
;      REGISTRATION NUMBER:  36,129
;      REFERENCE/DOCKET NUMBER:  CGC 1748/CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8614
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  537 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-472-028A-2

  Query Match             73.0%;  Score 2075;  DB 1;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 5.7e-198;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy     22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
          |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db      3 LSLLRPTTQSLLPSFSKPN-LRLNVYKPLRLRCSVAGGPTVGSSKIEGGGGTTITTDCVI 61

Qy     77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
          || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db     62 VGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITREENGFLWEEGPNSFQPSDP 121

Qy    133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
          |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db    122 MLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTDLPFFDLMSIGGKIRAGFGAL 181

Qy    193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
          |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||: 
Db    182 GIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGVYAGDPSKLSMKAAFGKVWKLEQN 241

Qy    253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
          ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db    242 GGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTVGSFRKGLRMLPEAISARLGSKVKLS 301

Qy    313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
          |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db    302 WKLSGITKLESGGYNLTYETPDGLVSVQSKSVVMTVPSHVASGLLRPLSESAANALSKLY 361

Qy    373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
          |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db    362 YPPVAAVSISYPKEAIRTECLIDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPPGRI 421

Qy    433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
          ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db    422 LLLNYIGGSTNTGILSKSEGELVEAVDRDLRKMLIKPNSTDPLKLGVRVWPQAIPQFLVG 481

Qy    493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
          | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db    482 HFDILDTAKSSLTSSGYEGLFLGGNYVAGVALGRCVEGAYETAIEVNNFMSRYAYK 537

RESULT   6
US-08-808-931-2
; Sequence 2, Application US/08808931
; Patent No. 5939602
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5939602artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,931
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  537 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-808-931-2

  Query Match             73.0%;  Score 2075;  DB 2;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 5.7e-198;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy       22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
            |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db        3 LSLLRPTTQSLLPSFSKPN-LRLNVYKPLRLRCSVAGGPTVGSSKIEGGGGTTITTDCVI 61

Qy       77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
            || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db       62 VGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITREENGFLWEEGPNSFQPSDP 121

Qy      133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
            |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db      122 MLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTDLPFFDLMSIGGKIRAGFGAL 181

Qy      193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
            |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||:
Db      182 GIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGVYAGDPSKLSMKAAFGKVWKLEQN 241

Qy      253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
            ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db      242 GGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTVGSFRKGLRMLPEAISARLGSKVKLS 301

Qy      313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
            |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db      302 WKLSGITKLESGGYNLTYETPDGLVSVQSKSVVMTVPSHVASGLLRPLSESAANALSKLY 361

Qy      373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
            |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db      362 YPPVAAVSISYPKEAIRTECLIDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPPGRI 421

Qy      433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
            ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db      422 LLLNYIGGSTNTGILSKSEGELVEAVDRDLRKMLIKPNSTDPLKLGVRVWPQAIPQFLVG 481

Qy      493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
            | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db      482 HFDILDTAKSSLTSSGYEGLFLGGNYVAGVALGRCVEGAYETAIEVNNFMSRYAYK 537


RESULT   7
US-08-808-323-2
; Sequence 2, Application US/08808323
; Patent No. 6018105
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Ward, Eric
;    TITLE OF INVENTION:  Promoters from Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase Genes
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6018105artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,323
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1846
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  537 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-808-323-2

  Query Match             73.0%;  Score 2075;  DB 3;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 5.7e-198;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;
```

```
Qy      22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
           |: |  |:     | || | : |  :   | |||||   || || ::||    :  ||||
Db       3 LSLLRPTTQSLLPSFSKPN-LRLNVYKPLRLRCSVAGGPTVGSSKIEGGGGTTITTDCVI 61

Qy      77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
           || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db      62 VGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITREENGFLWEEGPNSFQPSDP 121

Qy     133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
           |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db     122 MLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTDLPFFDLMSIGGKIRAGFGAL 181

Qy     193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
           |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||:
Db     182 GIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGVYAGDPSKLSMKAAFGKVWKLEQN 241

Qy     253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
           ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db     242 GGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTVGSFRKGLRMLPEAISARLGSKVKLS 301

Qy     313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
           |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db     302 WKLSGITKLESGGYNLTYETPDGLVSVQSKSVVMTVPSHVASGLLRPLSESAANALSKLY 361

Qy     373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
           |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db     362 YPPVAAVSISYPKEAIRTECLIDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPPGRI 421

Qy     433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
           ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db     422 LLLNYIGGSTNTGILSKSEGELVEAVDRDLRKMLIKPNSTDPLKLGVRVWPQAIPQFLVG 481

Qy     493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
           | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db     482 HFDILDTAKSSLTSSGYEGLFLGGNYVAGVALGRCVEGAYETAIEVNNFMSRYAYK 537

RESULT   8
US-09-050-603A-2
; Sequence 2, Application US/09050603A
; Patent No. 6023012
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6023012artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/050,603A
;      FILING DATE:  30-MAR-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  537 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-050-603A-2

  Query Match             73.0%;  Score 2075;  DB 3;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 5.7e-198;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy      22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
           |: |  |:     | || | : |  :   | |||||   || || ::||    :  ||||
Db       3 LSLLRPTTQSLLPSFSKPN-LRLNVYKPLRLRCSVAGGPTVGSSKIEGGGGTTITTDCVI 61

Qy      77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
           || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db      62 VGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITREENGFLWEEGPNSFQPSDP 121

Qy     133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
           |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db     122 MLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTDLPFFDLMSIGGKIRAGFGAL 181

Qy     193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
           |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||:
Db     182 GIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGVYAGDPSKLSMKAAFGKVWKLEQN 241

Qy     253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
           ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db     242 GGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTVGSFRKGLRMLPEAISARLGSKVKLS 301

Qy     313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
           |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db     302 WKLSGITKLESGGYNLTYETPDGLVSVQSKSVVMTVPSHVASGLLRPLSESAANALSKLY 361

Qy     373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
           |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db     362 YPPVAAVSISYPKEAIRTECLIDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPPGRI 421

Qy     433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
           ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db     422 LLLNYIGGSTNTGILSKSEGELVEAVDRDLRKMLIKPNSTDPLKLGVRVWPQAIPQFLVG 481

Qy     493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
           | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db     482 HFDILDTAKSSLTSSGYEGLFLGGNYVAGVALGRCVEGAYETAIEVNNFMSRYAYK 537

RESULT   9
US-09-102-420B-2
; Sequence 2, Application US/09102420B
; Patent No. 6084155
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
```

```
;    TITLE OF INVENTION:  OXIDASE ("PROTOX")
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6084155artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/102,420B
;      FILING DATE:  22-JUN-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/059,164
;      FILING DATE:  13-APR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/050,603
;      FILING DATE:  30-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/126,430
;      FILING DATE:  11-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/472,028
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847/CIP3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  537 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-102-420B-2


  Query Match             73.0%;  Score 2075;  DB 3;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 5.7e-198;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy         22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
              |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db          3 LSLLRPTTQSLLPSFSKPN-LRLNVYKPLRLRCSVAGGPTVGSSKIEGGGGTTITTDCVI 61

Qy         77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
              || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db         62 VGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITREENGFLWEEGPNSFQPSDP 121

Qy        133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
              |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db        122 MLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTDLPFFDLMSIGGKIRAGFGAL 181
```

```
Qy        193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
              |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||:
Db        182 GIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGVYAGDPSKLSMKAAFGKVWKLEQN 241

Qy        253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
              ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db        242 GGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTVGSFRKGLRMLPEAISARLGSKVKLS 301

Qy        313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
              |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db        302 WKLSGITKLESGGYNLTYETPDGLVSVQSKSVVMTVPSHVASGLLRPLSESAANALSKLY 361

Qy        373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
              |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db        362 YPPVAAVSISYPKEAIRTECLIDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPPGRI 421

Qy        433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
              ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db        422 LLLNYIGGSTNTGILSKSEGELVEAVDRDLRKMLIKPNSTDPLKLGVRVWPQAIPQFLVG 481

Qy        493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
              | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db        482 HFDILDTAKSSLTSSGYEGLFLGGNYVAGVALGRCVEGAYETAIEVNNFMSRYAYK 537

RESULT 10
US-09-071-296-2
; Sequence 2, Application US/09071296
; Patent No. 6177245
;  GENERAL INFORMATION:
;    APPLICANT:  Ward, Eric R
;    APPLICANT:  Volrath, Sandra
;    TITLE OF INVENTION:  Manipulation of Protoporphyrinogen
;    TITLE OF INVENTION:  Oxidase Enzyme Activity in Eukaryotic Organisms
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ciba-Geigy Corporation
;      STREET:  7 Skyline Drive
;      CITY:  Hawthorne
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10532
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,296
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/261,198
;      FILING DATE:  16-JUN-94
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Elmer, James Scott
;      REGISTRATION NUMBER:  36,129
;      REFERENCE/DOCKET NUMBER:  CGC 1748/CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8614
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  537 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-071-296-2


  Query Match             73.0%;  Score 2075;  DB 4;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 5.7e-198;
```

```
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy       22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
            |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db        3 LSLLRPTTQSLLPSFSKPN-LRLNVYKPLRLRCSVAGGPTVGSSKIEGGGGTTITTDCVI 61

Qy       77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
            || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db       62 VGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITREENGFLWEEGPNSFQPSDP 121

Qy      133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
            |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db      122 MLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTDLPFFDLMSIGGKIRAGFGAL 181

Qy      193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
            |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||:
Db      182 GIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGVYAGDPSKLSMKAAFGKVWKLEQN 241

Qy      253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
            ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db      242 GGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTVGSFRKGLRMLPEAISARLGSKVKLS 301

Qy      313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
            |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db      302 WKLSGITKLESGGYNLTYETPDGLVSVQSKSVVMTVPSHVASGLLRPLSESAANALSKLY 361

Qy      373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
            |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db      362 YPPVAAVSISYPKEAIRTECLIDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPPGRI 421

Qy      433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
            ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db      422 LLLNYIGGSTNTGILSKSEGELVEAVDRDLRKMLIKPNSTDPLKLGVRVWPQAIPQFLVG 481

Qy      493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
            | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db      482 HFDILDTAKSSLTSSGYEGLFLGGNYVAGVALGRCVEGAYETAIEVNNFMSRYAYK 537

RESULT  11
US-08-808-931-12
; Sequence 12, Application US/08808931
; Patent No. 5939602
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5939602artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,931
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  543 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-808-931-12

  Query Match             72.3%;  Score 2055;  DB 2;  Length 543;
  Best Local Similarity   78.1%;  Pred. No. 5.7e-196;
  Matches  392;  Conservative   44;  Mismatches   60;  Indels    6;  Gaps    3;

Qy       50 RCSVAKDYTV-PSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY--PNLMVTEARDRA 106
            |||:|:: |  |    |  |   :|||:|| |:||||||| ::  :   |::|||||||
Db       45 RCSIAEESTASPPKTRDSAP---VDCVVVGGGVSGLCIAQALATKHANANVVVTEARDRV 101

Qy      107 GGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPV 166
            ||||||:||||||||||||||||||||||| || ||||:||||||:|||||||  |||||
Db      102 GGNITTMERDGYLWEEGPNSFQPSDPMLTMVVDSGLKDELVLGDPDAPRFVLWNRKLRPV 161

Qy      167 PSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFC 226
            | ||||||||||||| ||:||||| :|:|| |||||||||:||||||| |||||||||||
Db      162 PGKLTDLPFFDLMSIGGKIRAGFGALGIRPPPPGHEESVEEFVRRNLGDEVFERLIEPFC 221

Qy      227 SGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQT 286
            |||| |||||||||||||||||||: ||||||||||||:||:   | |||||||||||||
Db      222 SGVYAGDPSKLSMKAAFGKVWKLEKNGGSIIGGTFKAIQERNGASKPPRDPRLPKPKGQT 281

Qy      287 VGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVM 346
            |||||||| |||||||||||:|:|||||||||:| : | | ||||||||||||| :::|:
Db      282 VGSFRKGLTMLPDAISARLGNKVKLSWKLSSISKLDSGEYSLTYETPEGVVSLQCKTVVL 341

Qy      347 TVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHP 406
            |:|||||| :||||| ||||||| |||||| ||:||||:|||| | |:||||||||||||
Db      342 TIPSYVASTLLRPLSAAAADALSKFYYPPVAAVSISYPKEAIRSECLIDGELKGFGQLHP 401

Qy      407 RTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKML 466
            |:|||||||||||||||||||| |||||||||||| |  |||||:|:||| |||||||:|
Db      402 RSQGVETLGTIYSSSLFPNRAPPGRVLLLNYIGGATNTGILSKTDSELVETVDRDLRKIL 461

Qy      467 IKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGR 526
            | | |||| |||||:|||||||||||||| |  |||:: : | |||||||||||||||||
Db      462 INPNAQDPFVVGVRLWPQAIPQFLVGHLDLLDVAKASIRNTGFEGLFLGGNYVSGVALGR 521

Qy      527 CVEGAYEVASEVTGFLSRYAYK 548
            |||||||||:||  ||:   ||
Db      522 CVEGAYEVAAEVNDFLTNRVYK 543

RESULT  12
US-08-808-323-12
; Sequence 12, Application US/08808323
; Patent No. 6018105
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Ward, Eric
;    TITLE OF INVENTION:  Promoters from Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase Genes
;    NUMBER OF SEQUENCES:  26
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6018105artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,323
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1846
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  543 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-808-323-12


  Query Match             72.3%;  Score 2055;  DB 3;  Length 543;
  Best Local Similarity   78.1%;  Pred. No. 5.7e-196;
  Matches  392;  Conservative   44;  Mismatches   60;  Indels    6;  Gaps    3;

Qy      50 RCSVAKDYTV-PSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY--PNLMVTEARDRA 106
           |||:|:: |  |    |  |   :|||:|| |:||||||| ::  :   |::||||||| 
Db      45 RCSIAEESTASPPKTRDSAP---VDCVVVGGGVSGLCIAQALATKHANANVVVTEARDRV 101

Qy     107 GGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPV 166
           ||||||:||||||||||||||||||||||| || ||||:||||||:|||||||  |||||
Db     102 GGNITTMERDGYLWEEGPNSFQPSDPMLTMVVDSGLKDELVLGDPDAPRFVLWNRKLRPV 161

Qy     167 PSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFC 226
           | ||||||||||||| ||:||||| :|:|| |||||||||:||||||| |||||||||||
Db     162 PGKLTDLPFFDLMSIGGKIRAGFGALGIRPPPPGHEESVEEFVRRNLGDEVFERLIEPFC 221

Qy     227 SGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQT 286
           |||| |||||||||||||||||||: ||||||||||||:||:   | |||||||||||||
Db     222 SGVYAGDPSKLSMKAAFGKVWKLEKNGGSIIGGTFKAIQERNGASKPPRDPRLPKPKGQT 281

Qy     287 VGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVM 346
           |||||||| |||||||||||:|:|||||||||:| : | | ||||||||||||| :::|:
Db     282 VGSFRKGLTMLPDAISARLGNKVKLSWKLSSISKLDSGEYSLTYETPEGVVSLQCKTVVL 341

Qy     347 TVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHP 406
           |:|||||| :||||| ||||||| |||||| ||:||||:|||| | |:||||||||||||
Db     342 TIPSYVASTLLRPLSAAAADALSKFYYPPVAAVSISYPKEAIRSECLIDGELKGFGQLHP 401

Qy     407 RTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKML 466
           |:|||||||||||||||||||| |||||||||||| |  |||||:|:||| |||||||:|
Db     402 RSQGVETLGTIYSSSLFPNRAPPGRVLLLNYIGGATNTGILSKTDSELVETVDRDLRKIL 461
```

```
Qy     467 IKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGR 526
           | | |||| |||||:|||||||||||||| |  |||:: : | |||||||||||||||||
Db     462 INPNAQDPFVVGVRLWPQAIPQFLVGHLDLLDVAKASIRNTGFEGLFLGGNYVSGVALGR 521

Qy     527 CVEGAYEVASEVTGFLSRYAYK 548
           |||||||||:||  ||:   ||
Db     522 CVEGAYEVAAEVNDFLTNRVYK 543


RESULT  13
US-09-050-603A-12
; Sequence 12, Application US/09050603A
; Patent No. 6023012
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6023012artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/050,603A
;      FILING DATE:  30-MAR-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  543 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-050-603A-12


  Query Match             72.3%;  Score 2055;  DB 3;  Length 543;
  Best Local Similarity   78.1%;  Pred. No. 5.7e-196;
  Matches  392;  Conservative   44;  Mismatches   60;  Indels    6;  Gaps    3;

Qy      50 RCSVAKDYTV-PSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY--PNLMVTEARDRA 106
           |||:|:: |  |    |  |   :|||:|| |:||||||| ::  :   |::||||||| 
Db      45 RCSIAEESTASPPKTRDSAP---VDCVVVGGGVSGLCIAQALATKHANANVVVTEARDRV 101
```

```
Qy    107 GGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPV 166
          ||||||:||||||||||||||||||||||| || ||||:||||||:|||||||  |||||
Db    102 GGNITTMERDGYLWEEGPNSFQPSDPMLTMVVDSGLKDELVLGDPDAPRFVLWNRKLRPV 161

Qy    167 PSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFC 226
          | ||||||||||||| ||:||||| :|:|| |||||||||:||||||| |||||||||||
Db    162 PGKLTDLPFFDLMSIGGKIRAGFGALGIRPPPPGHEESVEEFVRRNLGDEVFERLIEPFC 221

Qy    227 SGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQT 286
          |||| |||||||||||||||||||: ||||||||||||:||:   | |||||||||||||
Db    222 SGVYAGDPSKLSMKAAFGKVWKLEKNGGSIIGGTFKAIQERNGASKPPRDPRLPKPKGQT 281

Qy    287 VGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVM 346
          |||||||| |||||||||||:|:|||||||||:| : | | ||||||||||||| :::|:
Db    282 VGSFRKGLTMLPDAISARLGNKVKLSWKLSSISKLDSGEYSLTYETPEGVVSLQCKTVVL 341

Qy    347 TVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHP 406
          |:|||||| :||||| ||||||| |||||| ||:||||:|||| | |:||||||||||||
Db    342 TIPSYVASTLLRPLSAAAADALSKFYYPPVAAVSISYPKEAIRSECLIDGELKGFGQLHP 401

Qy    407 RTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKML 466
          |:|||||||||||||||||||| |||||||||||| |  |||||:|:||| |||||||:|
Db    402 RSQGVETLGTIYSSSLFPNRAPPGRVLLLNYIGGATNTGILSKTDSELVETVDRDLRKIL 461

Qy    467 IKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGR 526
          | | |||| |||||:|||||||||||||| |  |||:: : | |||||||||||||||||
Db    462 INPNAQDPFVVGVRLWPQAIPQFLVGHLDLLDVAKASIRNTGFEGLFLGGNYVSGVALGR 521

Qy    527 CVEGAYEVASEVTGFLSRYAYK 548
          |||||||||:||  ||:   ||
Db    522 CVEGAYEVAAEVNDFLTNRVYK 543

RESULT  14
US-09-102-420B-12
; Sequence 12, Application US/09102420B
; Patent No. 6084155
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
;    TITLE OF INVENTION:  OXIDASE ("PROTOX")
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6084155artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/102,420B
;      FILING DATE:  22-JUN-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/059,164
;      FILING DATE:  13-APR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/050,603
;      FILING DATE:  30-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/126,430
;      FILING DATE:  11-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/472,028
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847/CIP3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  543 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-102-420B-12

  Query Match             72.3%;  Score 2055;  DB 3;  Length 543;
  Best Local Similarity   78.1%;  Pred. No. 5.7e-196;
  Matches  392;  Conservative   44;  Mismatches   60;  Indels    6;  Gaps    3;

Qy     50 RCSVAKDYTV-PSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY--PNLMVTEARDRA 106
          |||:|:: |  |    |  |   :|||:|| |:||||||| ::  :   |::|||||||
Db     45 RCSIAEESTASPPKTRDSAP---VDCVVVGGGVSGLCIAQALATKHANANVVVTEARDRV 101

Qy    107 GGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPV 166
          ||||||:||||||||||||||||||||||| || ||||:||||||:|||||||  |||||
Db    102 GGNITTMERDGYLWEEGPNSFQPSDPMLTMVVDSGLKDELVLGDPDAPRFVLWNRKLRPV 161

Qy    167 PSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFC 226
          | ||||||||||||| ||:||||| :|:|| |||||||||:||||||| |||||||||||
Db    162 PGKLTDLPFFDLMSIGGKIRAGFGALGIRPPPPGHEESVEEFVRRNLGDEVFERLIEPFC 221

Qy    227 SGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQT 286
          |||| |||||||||||||||||||: ||||||||||||:||:   | |||||||||||||
Db    222 SGVYAGDPSKLSMKAAFGKVWKLEKNGGSIIGGTFKAIQERNGASKPPRDPRLPKPKGQT 281

Qy    287 VGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVM 346
          |||||||| |||||||||||:|:|||||||||:| : | | ||||||||||||| :::|:
Db    282 VGSFRKGLTMLPDAISARLGNKVKLSWKLSSISKLDSGEYSLTYETPEGVVSLQCKTVVL 341

Qy    347 TVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHP 406
          |:|||||| :||||| ||||||| |||||| ||:||||:|||| | |:||||||||||||
Db    342 TIPSYVASTLLRPLSAAAADALSKFYYPPVAAVSISYPKEAIRSECLIDGELKGFGQLHP 401

Qy    407 RTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKML 466
          |:|||||||||||||||||||| |||||||||||| |  |||||:|:||| |||||||:|
Db    402 RSQGVETLGTIYSSSLFPNRAPPGRVLLLNYIGGATNTGILSKTDSELVETVDRDLRKIL 461

Qy    467 IKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGR 526
          | | |||| |||||:|||||||||||||| |  |||:: : | |||||||||||||||||
Db    462 INPNAQDPFVVGVRLWPQAIPQFLVGHLDLLDVAKASIRNTGFEGLFLGGNYVSGVALGR 521

Qy    527 CVEGAYEVASEVTGFLSRYAYK 548
          |||||||||:||  ||:   ||
Db    522 CVEGAYEVAAEVNDFLTNRVYK 543

RESULT  15
US-08-808-931-20
; Sequence 20, Application US/08808931
```

```
; Patent No. 5939602
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5939602artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,931
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  536 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
US-08-808-931-20

  Query Match            70.6%;  Score 2006.5;  DB 2;  Length 536;
  Best Local Similarity  73.1%;  Pred. No. 3.8e-191;
  Matches  391;  Conservative   52;  Mismatches   83;  Indels    9;  Gaps    4;

Qy     22 LAFLNRTSFI-PFSSISKRNSVNCNGWRTRCSVAKDYTVPSSAVDGGPAAE---LDCVIV 77
          |: |    |: |||:   | |        ||||:    | || ::||   :    |||||
Db      3 LSLLRPQPFLSPFSNPFPR-SRPYKPLNLRCSVSGGSVVGSSTIEGGGGGKTVTADCVIV 61

Qy     78 GAGISGLCIAQVMSANYP----NLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDPM 133
          | ||||||||| :   :|    |:|||||:|| |||| | |  |:|||||||||||||||
Db     62 GGGISGLCIAQALVTKHPDAAKNVMVTEAKDRVGGNIITREEQGFLWEEGPNSFQPSDPM 121

Qy    134 LTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPIG 193
          ||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| ||
Db    122 LTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTDLPFFDLMSIGGKIRAGFGAIG 181

Qy    194 LRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEETG 253
          :|||||| |||||:||||||| ||||||||||||||| |||:|||||||||||||||| |
Db    182 IRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGVYAGDPAKLSMKAAFGKVWKLEENG 241

Qy    254 GSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLSW 313
          |||||| ||||: ::  ||  |||||||||||||||||||| |||:||||||| |:|:||
Db    242 GSIIGGAFKAIQAKNKAPKTTRDPRLPKPKGQTVGSFRKGLTMLPEAISARLGDKVKVSW 301

Qy    314 KLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFYY 373
          |||||||   | | ||||||||:|::||:|:||||||:|||::||||| :||:|||  ||
Db    302 KLSSITKLASGEYSLTYETPEGIVTVQSKSVVMTVPSHVASSLLRPLSDSAAEALSKLYY 361

Qy    374 PPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRVL 433
          ||| ||:||| :|||| | |:||||||||||||||| |||||||||||||||||| ||||
Db    362 PPVAAVSISYAKEAIRSECLIDGELKGFGQLHPRTQKVETLGTIYSSSLFPNRAPPGRVL 421

Qy    434 LLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGH 493
          |||||||| |  ||||:| :||| |||||||||||| : ||||:||::|||||||||:||
Db    422 LLNYIGGATNTGILSKSEGELVEAVDRDLRKMLIKPSSTDPLVLGVKLWPQAIPQFLIGH 481

Qy    494 LDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
          :| :  |||::: :| ||||||||||:||||||||||||| |::|  |:||||||
Db    482 IDLVDAAKASLSSSGHEGLFLGGNYVAGVALGRCVEGAYETATQVNDFMSRYAYK 536


Search completed: July  3, 2001, 10:31:54
Job time: 1913 sec
```

```
                         GenCore version 4.5
                Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          July  3, 2001, 10:33:07 ; Search time 34.09 Seconds
                                            (without alignments)
                                            550.660 Million cell updates/sec

Title:           US-09-508-418-2
Perfect score:   2844
Sequence:        1 MTTTPIANHPNIFTHQSSSS..........EGAYEVASEVTGFLSRYAYK 548

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                               SUMMARIES
                 %
Result          Query
   No.   Score  Match Length DB  ID                       Description
-------------------------------------------------------------------------
     1    2821   99.2    548  1  PPOC_TOBAC               O24163 nicotiana t
     2    2075   73.0    537  1  PPOC_ARATH               P55826 arabidopsis
     3     541   19.0    471  1  PPOX_MYXXA               P56601 myxococcus
     4     505   17.8    504  1  PPOM_TOBAC               O24164 nicotiana t
     5     485   17.1    470  1  PPOX_BACSU               P32397 bacillus su
     6   375.5   13.2    527  1  PPOX_PROFR               O32434 propionibac
     7   363.5   12.8    477  1  PPOX_HUMAN               P50336 homo sapien
     8   362.5   12.7    477  1  PPOX_MOUSE               P51175 mus musculu
     9   302.5   10.6    490  1  YAM7_SCHPO               Q10062 schizosacch
    10     221    7.8    426  1  PPOX_MYCTU               O53230 mycobacteri
    11   216.5    7.6    451  1  PPOX_MYCLE               Q50008 mycobacteri
    12     189    6.6    539  1  PPOX_YEAST               P40012 saccharomyc
    13     141    5.0    448  1  AOFH_MYCTU               O53320 mycobacteri
    14     139    4.9    570  1  CRTI_NARPS               Q40406 narcissus p
    15     134    4.7    522  1  AOF_ONCMY                P49253 oncorhynchu
    16   132.5    4.7    527  1  AOFA_BOVIN               P21398 bos taurus
    17     131    4.6    566  1  CRTI_ARATH               Q07356 arabidopsis
    18   125.5    4.4    417  1  Y4AB_RHISN               P55349 rhizobium s
    19   125.5    4.4    472  1  CRTI_SYNY3               P29273 synechocyst
    20   125.5    4.4    489  1  CBP1_CANAL               P31225 candida alb
    21   123.5    4.3    526  1  AOFA_RAT                 P21396 rattus norv
    22     123    4.3    582  1  CRTI_CAPAN               P80093 capsicum an
    23     122    4.3    471  1  Y782_SYNY3               Q55629 synechocyst
    24   119.5    4.2    478  1  PUO_MICRU                P40974 micrococcus
    25     118    4.1    571  1  CRTI_MAIZE               P49086 zea mays (m
    26   117.5    4.1    469  1  P49_STRLI                P06108 streptomyce
    27     116    4.1    583  1  CRTI_LYCES               P28554 lycopersico
    28   114.5    4.0   1106  1  ITA7_RAT                 Q63258 rattus norv
    29     114    4.0    531  1  ACHC_ACHFU               P35903 achatina fu
    30   112.5    4.0    570  1  CRTI_SOYBN               P28553 glycine max
    31    .111    3.9    508  1  CRTI_STRSE               P54971 streptomyce
    32   109.5    3.9    527  1  AOFA_HUMAN               P21397 homo sapien
    33   109.5    3.9    908  1  SRCA_RABIT               P13666 oryctolagus
```

```
    34     109    3.8   1181  1  ITA7_HUMAN               Q13683 homo sapien
    35   108.5    3.8   1062  1  CERU_MOUSE               Q61147 mus musculu
    36     108    3.8    504  1  A37C_DROME               P18487 drosophila
    37   107.5    3.8    516  1  ER13_ARATH               O65403 arabidopsis
    38     107    3.8    405  1  SAOX_CORS1               P40875 corynebacte
    39     107    3.8   1179  1  ITA7_MOUSE               Q61738 mus musculu
    40   106.5    3.7    630  1  FIG1_MOUSE               O09046 mus musculu
    41   105.5    3.7    872  1  IP3L_RAT                 P42335 rattus norv
    42     105    3.7    755  1  TR2M_AGRT3               P25017 agrobacteri
    43   104.5    3.7    466  1  AMID_PSEPU               O69768 pseudomonas
    44   104.5    3.7    517  1  ER12_ARATH               O65402 arabidopsis
    45   104.5    3.7    524  1  CRTI_RHOCA               P17054 rhodobacter


                               ALIGNMENTS


RESULT   1
PPOC_TOBAC
ID   PPOC_TOBAC     STANDARD;      PRT;   548 AA.
AC   O24163;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO I)
DE   (PROTOPORPHYRINOGEN IX OXIDASE ISOZYME I) (PPX I).
GN   PPXI.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. SR1;
RX   MEDLINE=97385200; PubMed=9238074;
RA   Lermontova I., Kruse E., Mock H.-P., Grimm B.;
RT   "Cloning and characterization of a plastidal and a mitochondrial
RT   isoform of tobacco protoporphyrinogen IX oxidase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:8895-8900(1997).
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
CC   -!- COFACTOR: FAD (BY SIMILARITY).
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS, AND
CC       CHLOROPHYLL SYNTHESIS.
CC   -!- SUBUNIT: HOMODIMER; CONTAINS ONE FAD PER HOMODIMER (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED IN EXPANDING PREMATURE LEAVES.
CC       DECREASED EXPRESSION IN OLDEST LEAVES. EXPRESSED AT VERY LOW LEVEL
CC       IN ROOTS.
CC   -!- INDUCTION: OSCILLATING EXPRESSION DURING DIURNAL GROWTH. MAXIMAL
CC       EXPRESSION IN THE DARK PERIOD.
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y13465; CAA73865.1; -.
KW   Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW   Transit peptide; Flavoprotein; FAD; Chloroplast;
KW   Chlorophyll biosynthesis.
FT   TRANSIT       1     50       CHLOROPLAST (POTENTIAL).
FT   CHAIN        51    548       PROTOPORPHYRINOGEN OXIDASE.
FT   NP_BIND      78     83       FAD (POTENTIAL).
SQ   SEQUENCE   548 AA;  59230 MW;  66892E78FB8A3E30 CRC64;
```

```
  Query Match             99.2%;  Score 2821;  DB 1;  Length 548;
  Best Local Similarity   99.3%;  Pred. No. 1.4e-205;
  Matches  544;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60

Qy      61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120

Qy     121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLAFFDLMS 180

Qy     181 IPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMK 240
           |||||||||| ||||||||||||||||||||||||||||||||||||||| |||||||||
Db     181 IPGKLRAGFGAIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYAGDPSKLSMK 240

Qy     241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300

Qy     301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360

Qy     361 SVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420
           |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db     361 SVAAADALSNFYYPPVGAVTITYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420

Qy     421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480

Qy     481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540

Qy     541 FLSRYAYK 548
           ||||||||
Db     541 FLSRYAYK 548


RESULT   2
PPOC_ARATH
ID   PPOC_ARATH     STANDARD;      PRT;   537 AA.
AC   P55826;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO).
GN   PPOX OR AT4G01690 OR T15B16.13.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97136707; PubMed=8982084;
RA   Narita S., Tanaka R., Ito T., Okada K., Taketani S., Inokuchi H.;
RT   "Molecular cloning and characterization of a cDNA that encodes
RT   protoporphyrinogen oxidase of Arabidopsis thaliana.";
RL   Gene 182:169-175(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sehkon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaugnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
CC   -!- COFACTOR: FAD (BY SIMILARITY).
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST (POTENTIAL).
CC   -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE LEAVES AND AT
CC       LOW LEVELS IN THE ROOTS AND FLORAL BUDS.
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D83139; BAA11820.1; -.
DR   EMBL; AL161492; CAB77739.1; -.
KW   Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW   Flavoprotein; FAD; Chloroplast; Transit peptide.
FT   TRANSIT       1      ?       CHLOROPLAST (POTENTIAL).
FT   CHAIN         ?    537       PROTOPORPHYRINOGEN OXIDASE.
FT   NP_BIND      63     68       FAD (POTENTIAL).
SQ   SEQUENCE   537 AA;  57695 MW;  C125B0EE1D3187D5 CRC64;
```

```
  Query Match             73.0%;  Score 2075;  DB 1;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 3.1e-149;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy      22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
           |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db       3 LSLLRPTTQSLLPSFSKPN-LRLNVYKPLRLRCSVAGGPTVGSSKIEGGGGTTITTDCVI 61

Qy      77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
           || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db      62 VGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITREENGFLWEEGPNSFQPSDP 121

Qy     133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
           |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db     122 MLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTDLPFFDLMSIGGKIRAGFGAL 181

Qy     193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
           |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||:
Db     182 GIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGVYAGDPSKLSMKAAFGKVWKLEQN 241

Qy     253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
           ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db     242 GGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTVGSFRKGLRMLPEAISARLGSKVKLS 301

Qy     313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
           |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db     302 WKLSGITKLESGGYNLTYETPDGLVSVQSKSVVMTVPSHVASGLLRPLSESAANALSKLY 361

Qy     373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
           |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db     362 YPPVAAVSISYPKEAIRTECLIDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPPGRI 421

Qy     433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
           ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db     422 LLLNYIGGSTNTGILSKSEGELVEAVDRDLRKMLIKPNSTDPLKLGVRVWPQAIPQFLVG 481

Qy     493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
           | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db     482 HFDILDTAKSSLTSSGYEGLFLGGNYVAGVALGRCVEGAYETAIEVNNFMSRYAYK 537

RESULT   3
PPOX_MYXXA
ID   PPOX_MYXXA     STANDARD;      PRT;   471 AA.
AC   P56601;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
GN   HEMY.
OS   Myxococcus xanthus.
OC   Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC   Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX   NCBI_TaxID=34;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92192453; PubMed=1372277;
RA   Li S., Lee B.U., Shimkets L.J.;
RT   "csgA expression entrains Myxococcus xanthus development.";
RL   Genes Dev. 6:401-410(1992).
RN   [2]
RP   REVISION, AND CHARACTERIZATION.
RX   MEDLINE=96224079; PubMed=8621504;
RA   Dailey H.A., Dailey T.A.;
RT   "Protoporphyrinogen oxidase of Myxococcus xanthus. Expression,
RT   purification, and characterization of the cloned enzyme.";
RL   J. Biol. Chem. 271:8714-8718(1996).
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX. MAY HAVE AN ESSENTIAL EXTRACELLULAR
CC       FUNCTION DURING DEVELOPMENT POSSIBLY AS A PHEROMONE.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
```

```
CC   -!- COFACTOR: CONTAINS ONE FAD PER HOMODIMER.
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC       FRAMESHIFT AT POSITION 127.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M73709; -; NOT_ANNOTATED_CDS.
DR   EMBL; AF098938; AAD13609.1; -.
KW   Porphyrin biosynthesis; Oxidoreductase; Flavoprotein; FAD;
KW   Heme biosynthesis.
FT   NP_BIND      16     21       FAD (POTENTIAL).
SQ   SEQUENCE   471 AA;  49387 MW;  2E9909D464F81515 CRC64;

  Query Match             19.0%;  Score 541;  DB 1;  Length 471;
  Best Local Similarity   31.7%;  Pred. No. 1.9e-33;
  Matches  150;  Conservative   84;  Mismatches  207;  Indels   32;  Gaps   12;

Qy      72 LDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSD 131
           ::  :|| ||||| :|  : :   : :: |:  | || : |    ||| |:|||||   :
Db      10 MNVAVVGGGISGLAVAHHLRSRGTDAVLLESSARLGGAVGTHALAGYLVEQGPNSFLDRE 69

Qy     132 PML-TMAVDCGLKDDLVLGDPNAP-RFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGF 189
           |    :|    |:  :   || |  |:|  :|:|| ||:        |:: :  :||
Db      70 PATRALAAALNLEGRIRAADPAAKRRYVYTRGRLRSVPASPPAFLASDILPLGARLRVA- 128

Qy     190 GPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKL 249
           | :  | :| | :||:  | ||:||    : |::   :|:| ||  :||: | |  : |:
Db     129 GELFSRRAPEGVDESLAAFGRRHLGHRATQVLLDAVQTGIYAGDVEQLSVAATFPMLVKM 188

Qy     250 EETGGSIIGGTFKAIKERSSTPKAPRDPRLP---KPK-GQTVGSFRKGLRMLPDAISARL 305
           |    |:| |  :|        || |   ||    ||    : :|  ||::| ||::| |
Db     189 EREHRSLILGAIRA-------QKAQRQAALPAGTAPKLSGALSTFDGGLQVLIDALAASL 241

Qy     306 GSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAA 365
           |    :  ::  : : | ||: |  |       |    :|:  |::  : :||||  | |
Db     242 GDAAHVGARVEGLAR-EDGGWRLIIEEHGRRAELSVAQVVLAAPAHATAKLLRPLDDALA 300

Qy     366 DALSNFYYPPVGAVTISYPQEAIRDERLVDGEL---KGFGQLHPRTQGVETLGTIYSSSL 422
             ::   | |:  | : :            | |    ||| | |  :    || |::|:
Db     301 ALVAGIAYAPIAVVHLGFDA----------GTLPAPDGFGFLVPAEEQRRMLGAIHASTT 350

Qy     423 FPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVW 482
           || ||  ||||    :|||: | :: : |  |  :   :| | |    |: |    |  |
Db     351 FPFRAEGGRVLYSCMVGGARQPGLVEQDEDALAALAREEL-KALAGVTAR-PSFTRVFRW 408

Qy     483 PQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVA 535
           |  |||: :|||: ::   ||:    | || | ||   || |  |:  | ::|
Db     409 PLGIPQYNLGHLERVAAIDAAL--QRLPGLHLIGNAYKGVGLNDCIRNAAQLA 459

RESULT   4
PPOM_TOBAC
ID   PPOM_TOBAC     STANDARD;      PRT;   504 AA.
AC   O24164;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE, MITOCHONDRIAL (EC 1.3.3.4) (PPO II)
DE   (PROTOPORPHYRINOGEN IX OXIDASE ISOZYME II) (PPX II) (PX-2).
GN   PPXII OR PPOX2.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
```

```
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. SR1;
RX   MEDLINE=97385200; PubMed=9238074;
RA   Lermontova I., Kruse E., Mock H.-P., Grimm B.;
RT   "Cloning and characterization of a plastidal and a mitochondrial
RT   isoform of tobacco protoporphyrinogen IX oxidase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:8895-8900(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. Samsun NN;
RX   PubMed=9808719;
RA   Watanabe N., Che F., Iwano M., Takayama S., Nakano T., Yoshida S.,
RA   Isogai A.;
RT   "Molecular characterization of photomixotrophic cultured tobacco cells
RT   resistant to protoporphyrinogen oxidase-inhibiting herbicides.";
RL   Plant Physiol. 118:751-758(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. SR1;
RA   Horikoshi M., Mametsuka K., Hirooka T.;
RT   "The molecular basis of photobleaching herbicide resistance in
RT   Tobacco.";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX.
CC   -!- FUNCTION: PROVIDES PRECURSOR FOR THE MITOCHONDRIAL AND PLASTIDIC
CC       HEME SYNTHESIS AND THE PREDOMINANT CHLOROPHYLL SYNTHESIS IN
CC       PLASTIDS.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) =
CC       PROTOPORPHYRIN-IX + H(2)O(2).
CC   -!- COFACTOR: FAD (BY SIMILARITY).
CC   -!- ENZYME REGULATION: INHIBITED BY THE HERBICIDE ACIFLUORFEN.
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRION.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED IN EXPANDING PREMATURE LEAVES.
CC       DECREASED EXPRESSION IN OLDEST LEAVES. ALSO DETECTED IN ROOTS.
CC   -!- INDUCTION: OSCILLATING EXPRESSION DURING DIURNAL GROWTH. MAXIMAL
CC       EXPRESSION IN THE DARK PERIOD.
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y13466; CAA73866.1; -.
DR   EMBL; AB020500; BAA34712.1; -.
DR   EMBL; AF044129; AAD02291.1; -.
DR   InterPro; IPR002937; -.
DR   Pfam; PF01593; Amino_oxidase; 1.
KW   Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW   Flavoprotein; FAD; Mitochondrion.
FT   NP_BIND      20     25       FAD (POTENTIAL).
SQ   SEQUENCE   504 AA;  55407 MW;  B85B55EC881DC00A CRC64;


  Query Match             17.8%;  Score 505;  DB 1;  Length 504;
  Best Local Similarity   28.8%;  Pred. No. 1.1e-30;
  Matches  149;  Conservative  100;  Mismatches  208;  Indels   60;  Gaps   17;

Qy         60 PSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYL 119
              ||:  |   :|:    ::|||:|||  |  :  :  |: | ||  :||| : :| :|| :
Db          3 PSAGEDKHSSAK-RVAVIGAGVSGLAAAYKLKIHGLNVTVFEAEGKAGGKLRSVSQDGLI 61

Qy        120 WEEGPNSFQPSDPMLTMAVD-CGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDL 178
              |:|| |:   |:  :|  :|  ||::          |::   |    :||   ||   :
Db         62 WDEGANTMTESEGDVTFLIDSLGLREKQQFPLSQNKRYIARNGTPVLLPSNPIDLIKSNF 121

Qy        179 MSIPGKLRAGFGPI-----GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGD 233
              :|   ||:    ||      |      | |||  | :|: | || : ||:|| :|   ||
Db        122 LSTGSKLQMLLEPILWKNKKLSQVSDSH-ESVSGFFQRHFGKEVVDYLIDPFVAGTCGGD 180

Qy        234 PSKLSMKAAFGKVWKLEETGGSIIGGTFKA-IKERSSTPKAPRDPRLPKPKGQTVGSFRK 292
              |  |||  :| ::| ||:  ||:| |  :: :  ::   : |      | : :   ||
Db        181 PDSLSMHHSFPELWNLEKRFGSVILGAIRSKLSPKNEKKQGPPKTSANKKRQRGSFSFLG 240

Qy        293 GLRMLPDAI-------SARLGSK-LKL-----------SWKLSSITKSEKGGYHLTYETP 333
              |:: | |||         || |: |:|           || : | :  ::     :::
Db        241 GMQTLTDAICKDLREDELRLNSRVLELSCSCTEDSAIDSWSIISASPHKRQSEEESFD-- 298

Qy        334 EGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY----YPPVGAVTISYPQEAIR 389
                       :::|| |     ::   ::      | ||     | |:  |  :: :| ::
Db        299 ---------AVIMTAPLCDVKSM--KIAKRGNPFLLNFIPEVDYVPLSVVITTFKRENVK 347

Qy        390 DERLVDGELKGFGQLHP---RTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEI 446
                      |:||| | |   :  |::||||::|| :||:|||    |   ::||::| |:
Db        348 ------YPLEGFGVLVPSKEQQHGLKTLGTLFSSMMFPDRAPNNVYLYTTFVGGSRNREL 401

Qy        447 LSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGH-LDTLSTAKAAMN 505
                 : ::| |:|  ||:::|      :|  |    | :| |  | ||  |::  |   |
Db        402 AKASRTELKEIVTSDLKQLL--GAEGEPTYVNHLYWSKAFP--LYGHNYDSVLDAIDKME 457

Qy        506 DNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFL 542
               | | |||  ||:  |:::|: :      |  |  :|
Db        458 KN-LPGLFYAGNHRGGLSVGKALSSGCNAADLVISYL 493


RESULT   5
PPOX_BACSU
ID   PPOX_BACSU     STANDARD;      PRT;   470 AA.
AC   P32397;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
GN   HEMY OR HEMG.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93094140; PubMed=1459957;
RA   Hansson M., Hederstedt L.;
RT   "Cloning and characterization of the Bacillus subtilis hemEHY gene
RT   cluster, which encodes protoheme IX biosynthetic enzymes.";
RL   J. Bacteriol. 174:8081-8093(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   CHARACTERIZATION.
RX   MEDLINE=94117488; PubMed=8288631;
RA   Dailey T.A., Meissner P., Dailey H.A.;
RT   "Expression of a cloned protoporphyrinogen oxidase.";
RL   J. Biol. Chem. 269:813-815(1994).
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX. ALSO OXIDIZES THE PATHWAY
CC       INTERMEDIATE COPROPORPHYRINOGEN III.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
CC   -!- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
CC   -!- PATHWAY: PENULTIMATE STEP IN PORPHYRIN BIOSYNTHESIS.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M97208; AAA22519.1; -.
DR   EMBL; Y14083; CAA74520.1; -.
DR   EMBL; Z99109; CAB12854.1; -.
DR   PIR; D47045; D47045.
DR   HSSP; P54725; 1UBA.
DR   SubtiList; BG10431; hemY.
KW   Porphyrin biosynthesis; Oxidoreductase; Flavoprotein; FAD.
FT   NP_BIND      12     17       FAD (POTENTIAL).
SQ   SEQUENCE   470 AA;  51203 MW;  95CC4E5847686D2E CRC64;

  Query Match             17.1%;  Score 485;  DB 1;  Length 470;
  Best Local Similarity   28.7%;  Pred. No. 3.2e-29;
  Matches  139;  Conservative   98;  Mismatches  206;  Indels   42;  Gaps   15;

Qy    75 VIVGAGISGLCIA-----QVMSANYP-NLMVTEARDRAGGNITTVERDGYLWEEGPNSF- 127
         ||:| ||:||  |     ::   | |  | : ||  | || | ||::|||: | ||:||
Db     9 VIIGGGITGLAAAFYMEKEIKEKNLPLELTLVEASPRVGGKIQTVKKDGYIIERGPDSFL 68

Qy   128 --QPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLW--------KGKLRPVPSKLTDLPFFD 177
           : | | |    | || : |:: :     :||         || :  :|:|:
Db    69 ERKKSAPQLVK--DLGL-EHLLVNNATGQSYVLVNRTLHPMPKGAVMGIPTKIAPFVSTG 125

Qy   178 LMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKL 237
         | |: || ||    | |  |    ::|: :| || :| || | ||||  ||:| ||  ||
Db   126 LFSLSGKARAAMDFI-LPASKTKDDQSLGEFFRRRVGDEVVENLIEPLLSGIYAGDIDKL 184

Qy   238 SMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRML 297
         |: : | : :: |:   |:| |       : : |:        | :||   :   ||: |
Db   185 SLMSTFPQFYQTEQKHRSLILGM------KKTRPQGSGQQLTAKKQGQ-FQTLSTGLQTL 237

Qy   298 PDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNIL 357
          : |  :|  ||   :| : :||    |   : |   | |:|:  |:::| |   |: :|
Db   238 VEEIEKQL--KLTKVYKGTKVTKLSHSGSCYSLELDNG-VTLDADSVIVTAPHKAAAGML 294

Qy   358 RPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTI 417
           |   |   | | :   |  | : :|: ::: |       :| | :  |
Db   295 SEL--PAISHLKNMHSTSVANVALGFPEGSVQMEH------EGTGFVISRNSDFAITACT 346

Qy   418 YSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVV 477
         :::  :|: ||:|: ||  |:| | :  |:  ::: :: :|  ||:|::      :| :
Db   347 WTNKKWPHAAPEGKTLLRAYVGKAGDESIVDLSDNDIINIVLEDLKKVM--NINGEPEMT 404

Qy   478 GVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASE 537
          |  | :::||: |||   :   : |:  :   |::: |    || :  |::      |:
Db   405 CVTRWHESMPQYHVGHKQRIKELREAL-ASAYPGVYMTGASFEGVGIPDCIDQGKAAVSD 463

Qy   538 VTGFL 542
            :|
Db   464 ALTYL 468

RESULT   6
PPOX_PROFR
ID   PPOX_PROFR     STANDARD;      PRT;   527 AA.
AC   O32434;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
GN   HEMY.
OS   Propionibacterium freudenreichii shermanii.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Propionibacterineae; Propionibacteriaceae;
OC   Propionibacterium.
```

```
OX   NCBI_TaxID=1752;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IFO 12424;
RX   MEDLINE=97306686; PubMed=9163953;
RA   Hashimoto Y., Yamashita Y., Murooka Y.;
RT   "The Propionibacterium freudenreichii hemYHBXRL gene cluster, which
RT   encodes enzymes and a regulator involved in the biosynthetic pathway
RT   from glutamate to protoheme.";
RL   Appl. Microbiol. Biotechnol. 47:385-392(1997).
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
CC   -!- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; D85417; BAA21909.1; -.
DR   InterPro; IPR000171; -.
DR   Pfam; PF02032; Phytoene_dh; 1.
KW   Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW   Flavoprotein; FAD.
FT   NP_BIND      33     38       FAD (POTENTIAL).
SQ   SEQUENCE   527 AA;  55267 MW;  579CB0A3318BF938 CRC64;

  Query Match             13.2%;  Score 375.5;  DB 1;  Length 527;
  Best Local Similarity   27.2%;  Pred. No. 7.1e-21;
  Matches  144;  Conservative   94;  Mismatches  210;  Indels   81;  Gaps   22;

Qy    58 TVPSSAVDG----GP-AAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITT 112
         | |:  | |    || |:    |:|| ||:||  |    |    : | |: | || : |
Db     8 TTPTPTVSGTDAPGPDASHCHLVVVGGGITGLAAAWQGMARGARVSVVESDDHFGGKVVT 67

Qy   113 VERDGYLWEEGPNSFQPSDP-MLTMAVDCGLKDDLVLGDPNAPRFV--LWKGKLRPVPSK 169
           |||:| |:||:||    |  | :  : || | ::   |   | |  | :|||||:|:
Db    68 DRRDGFLVEQGPDSFVAYRPAALKLIEELGLSDQVIA--PGGGRRVSLLSRGKLRPMPAG 125

Qy   170 LTDL------PFF--DLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERL 221
         :  :      ||    ::| | |:|||   : :    | |: ::  |:|: ||  :  |
Db   126 MGMVLPTRMWPFVTTTVLSWPDKIRAGLDLV-IPRRLPDHDVAIGAFLRQRLGDGIVRRF 184

Qy   222 IEPFCSGVYVGDPSKLSMKAAFGKVWKLEETG-----GSIIGG--TFKAIKERSSTPKAP 274
          :|   |:|     :||: |    :   |         |: ||  : :| ::|::   |
Db   185 ADPMVGGIYGAGIDELSLDAVLPSLRDNERDHRSLMVASLAGGRASRRAARQRAAQNNAQ 244

Qy   275 RDPRLPKPKGQ-----TVG-------SFRKGLRMLPDAISARL---GSKLKLSWKLSSIT 319
         ::       ||     | |       : | ||  | ||:  :|   | :| ::   :|:
Db   245 QNSSHQNSTGQNNSAGTRGPAASPFRTLRGGLGQLIDALVDQLRAGGVELLVN---TSVD 301

Qy   320 KSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAV 379
            : | ||:    :| | | : ::|:      :: :|||   ||| ||:         |
Db   302 LLGRDGVHLS----DGRV-LPADAVVLAGGVASSARLLRPQLPAAARALAQIPLASTTIV 356

Qy   380 TISYPQEAIRDERLVDGELKGFGQLHPRTQG------VETLGTIYSSSLFPNRAPKGRVL 433
         ::::|  |                : | :||          |   ||  |  ||| | ||
Db   357 SLAWPVSAF--------------DVAPDSQGWLEADAGPVSGLTASSIKFAGRAPDGSVL 402

Qy   434 LLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGH 493
         :  ::   : | :    : :|:  |   :| :|      :| :  :  | : :|:: |||
Db   403 MRVFVPDKRGP-LTDAPDDELLSAVIDHVRPLL--GVHGEPGLTQITRWHKVMPKYTVGH 459
```

```
Qy    494 LDTLSTAKAAMNDNGL----EGLFLGGNYVSGVALGRCVEGAYEVASEV 538
          |:     :||: |: |        : |: : || |  |:  |   | ||
Db    460 LE-----RAAVVDSTLAEQRPTWAVAGSALHGVGLPDCISDARHSADEV 503


RESULT   7
PPOX_HUMAN
ID   PPOX_HUMAN     STANDARD;      PRT;   477 AA.
AC   P50336;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
GN   PPOX.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=95229621; PubMed=7713909;
RA   Nishimura K., Taketani S., Inokuchi H.;
RT   "Cloning of a human cDNA for protoporphyrinogen oxidase by
RT   complementation in vivo of a hemG mutant of Escherichia coli.";
RL   J. Biol. Chem. 270:8076-8080(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=96367087; PubMed=8771201;
RA   Dailey T.A., Dailey H.A.;
RT   "Human protoporphyrinogen oxidase: expression, purification, and
RT   characterization of the cloned enzyme.";
RL   Protein Sci. 5:98-105(1996).
RN   [3]
RP   VARIANT VP ARG-232, AND VARIANT HIS-304.
RX   MEDLINE=97005368; PubMed=8852667;
RA   Deybach J.-C., Puy H., Robreau A.-M., Lamoril J., da Silva V.,
RA   Grandchamp B., Nordmann Y.;
RT   "Mutations in the protoporphyrinogen oxidase gene in patients with
RT   variegate porphyria.";
RL   Hum. Mol. Genet. 5:407-410(1996).
RN   [4]
RP   VARIANT VP TRP-59, AND VARIANT CYS-168.
RX   MEDLINE=96241580; PubMed=8673113;
RA   Meissner P.N., Dailey T.A., Hift R.J., Ziman M., Corrigall A.V.,
RA   Roberts A.G., Meissner D.M., Kirsch R.E., Dailey H.A.;
RT   "A R59W mutation in human protoporphyrinogen oxidase results in
RT   decreased enzyme activity and is prevalent in South Africans with
RT   variegate porphyria.";
RL   Nat. Genet. 13:95-97(1996).
RN   [5]
RP   VARIANT VP CYS-152.
RX   MEDLINE=98434271; PubMed=9763307;
RA   Frank J., Poh-Fitzpatrick M.B., King L.E. Jr., Christiano A.M.;
RT   "The genetic basis of 'Scarsdale Gourmet Diet' variegate porphyria: a
RT   missense mutation in the protoporphyrinogen oxidase gene.";
RL   Arch. Dermatol. Res. 290:441-445(1998).
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
CC   -!- COFACTOR: CONTAINS ONE FAD PER HOMODIMER.
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE
CC       WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG,
CC       LIVER, SKELETAL MUSCLE, KIDNEY, AND PANCREAS.
CC   -!- DISEASE: DEFECTS IN PPOX ARE THE CAUSE OF PORPHYRIA VARIEGATA
CC       (VP), A DISEASE CHARACTERIZED BY SKIN HYPERPIGMENTATION AND HAIR
CC       HYPERTRICHOSIS, ASSOCIATED WITH ACUTE ATTACKS, LIKE THOSE OF ACUTE
CC       INTERMITTENT PORPHYRIA.
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D38537; BAA07538.1; -.
DR   EMBL; U26446; AAA67690.1; -.
DR   MIM; 600923; -.
DR   MIM; 176200; -.
KW   Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW   Flavoprotein; FAD; Mitochondrion; Disease mutation; Polymorphism.
FT   NP_BIND       9     14       FAD (POTENTIAL).
FT   VARIANT      59     59       R -> W (IN VP).
FT                                /FTId=VAR_003686.
FT   VARIANT     152    152       R -> C (IN VP).
FT                                /FTId=VAR_003687.
FT   VARIANT     168    168       R -> C.
FT                                /FTId=VAR_003688.
FT   VARIANT     232    232       G -> R (IN VP).
FT                                /FTId=VAR_003689.
FT   VARIANT     304    304       R -> H.
FT                                /FTId=VAR_003690.
SQ   SEQUENCE   477 AA;  50765 MW;  2444DEAC2E6C33EE CRC64;

  Query Match             12.8%;  Score 363.5;  DB 1;  Length 477;
  Best Local Similarity   25.5%;  Pred. No. 5e-20;
  Matches  138;  Conservative   83;  Mismatches  170;  Indels  151;  Gaps   23;

Qy     75 VIVGAGISGLCIAQVMSANY--------PNLMVTEARDRAGGNITTVE-RDGYLWEEGPN 125
          |::| |||||      :|:|        | ::: |: :| || | :|   :| ::| ||
Db      6 VVLGGGISGL------AASYHLSRAPCPPKVVLVESSERLGGWIRSVRGPNGAIFELGPR 59

Qy    126 SFQPSDPM----LTMAVDCGLKDDL--VLGDPNAP--RFVLWKGKLRPVPSKLTDLPFFD 177
            :|:  :    | :  : ||  ::  | ||  |   ||:   | |  :|: |  |
Db     60 GIRPAGALGARTLLLVSELGLDSEVLPVRGDHPAAQNRFLYVGGALHALPTGLRGL---- 115

Qy    178 LMSIPGKLRAGFGPIGLRPSPP-----------------GHE--ESVEQFVRRNLGGEVF 218
                          ||||||                 | |  |:|  | :| || ||
Db    116 ----------------LRPSPPFSKPLFWAGLRELTKPRGKEPDETVHSFAQRRLGPEVA 159

Qy    219 ERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPR 278
             ::  | ||: |:  :||::: |  ::: |:|  ||: |                  |
Db    160 SLAMDSLCRGVFAGNSRELSIRSCFPSLFQAEQTHRSILLGLLLGA------------GR 207

Qy    279 LPKPKGQTV----------GSFRKGLRMLPDAISARLGSK------------LKLS---- 312
           |:|    :           | | || ||| |:   | |:            | |
Db    208 TPQPDSALIRQALAERWSQWSLRGGLEMLPQALETHLTSRGVSVLRGQPVCGLSLQAEGR 267

Qy    313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
          ||: |:  |                ||::  ::  :|: | | :|   :   | |||
Db    268 WKV-SLRDS----------------SLEADHVISAIPASVLSELLPAEAAPLARALSAIT 310

Qy    373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNR--APKG 430
             |  | : |    :         ::||| | | ::    || :| |  || :  :| |
Db    311 AVSVAVVNLQYQGAHL--------PVQGFGHLVPSSEDPGVLGIVYDSVAFPEQDGSPPG 362

Qy    431 -RVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLV--VGVRVWP---- 483
           || ::  :||         :  | :|     | : | : :||:     :|::  |
Db    363 LRVTVM--LGG---------SWLQTLEASGCVLSQELFQQRAQEAAATQLGLKEMPSHCL 411

Qy    484 -----QAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEV 538
                 |||: :||   | :|:  :  : |     | :|  |||:  |:|  :  |  |
Db    412 VHLHKNCIPQYTLGHWQKLESARQFLTAHRLPLTLAGASY-EGVAVNDCIESGRQAAVSV 470

Qy    539 TG 540
           |
Db    471 LG 472
```

```
RESULT   8
PPOX_MOUSE
ID   PPOX_MOUSE     STANDARD;      PRT;   477 AA.
AC   P51175; P97344;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
GN   PPOX.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96132666; PubMed=8554330;
RA   Dailey T.A., Dailey H.A., Meissner P., Prasad A.R.;
RT   "Cloning, sequence, and expression of mouse protoporphyrinogen
RT   oxidase.";
RL   Arch. Biochem. Biophys. 324:379-384(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Erythroleukemia;
RX   MEDLINE=95331315; PubMed=7607249;
RA   Taketani S., Yoshinaga T., Furukawa T., Kohno H., Tokunaga R.,
RA   Nishimura K., Inokuchi H.;
RT   "Induction of terminal enzymes for heme biosynthesis during
RT   differentiation of mouse erythroleukemia cells.";
RL   Eur. J. Biochem. 230:760-765(1995).
RN   [3]
RP   SUBCELLULAR LOCATION.
RC   TISSUE=Liver;
RX   MEDLINE=88153682; PubMed=3346226;
RA   Ferreira G.C., Andrew T.L., Karr S.W., Dailey H.A.;
RT   "Organization of the terminal two enzymes of the heme biosynthetic
RT   pathway. Orientation of protoporphyrinogen oxidase and evidence for a
RT   membrane complex.";
RL   J. Biol. Chem. 263:3835-3839(1988).
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
CC   -!- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE
CC       WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE.
CC   -!- INDUCTION: DURING ERYTHROID DIFFERENTIATION.
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U25114; AAA96003.1; -.
DR   EMBL; D45185; BAA08126.1; -.
DR   MGD; MGI:104968; Ppox.
KW   Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW   Flavoprotein; FAD; Mitochondrion.
FT   NP_BIND       9     14       FAD (POTENTIAL).
FT   CONFLICT     64     64       A -> T (IN REF. 2).
FT   CONFLICT     66     66       A -> P (IN REF. 2).
FT   CONFLICT    108    108       L -> S (IN REF. 2).
FT   CONFLICT    427    427       W -> C (IN REF. 2).
SQ   SEQUENCE   477 AA;  50870 MW;  8CFB48120728DE6F CRC64;

  Query Match             12.7%;  Score 362.5;  DB 1;  Length 477;
  Best Local Similarity   24.8%;  Pred. No. 5.9e-20;
  Matches  133;  Conservative   76;  Mismatches  188;  Indels  139;  Gaps   19;

Qy       75 VIVGAGISGLCIAQ--VMSANYPNLMVTEARDRAGGNITTVE-RDGYLWEEGPNSFQPSD 131
            :::| |||||  :   :   : | ::: |   | || | ::   || ::| ||   :|:
Db        6 IVLGGGISGLAASYHLIRGPSPPKVILVEGSKRLGGWIRSIRGSDGAIFELGPRGIRPAG 65

Qy      132 PM----LTMAVDCGLKDDL--VLGDPNAP--RFVLWKGKLRPVPSKLTDLPFFDLMSIPG 183
             :    | :  : ||: ::  | ||  |   ||:   | | |:|| |  |
Db       66 ALGARTLLLVSELGLESEVLPVRGDHPAAQNRFLYVGGTLHPLPSGLRGL---------- 115

Qy      184 KLRAGFGPIGLRPSPP-----------------GHE--ESVEQFVRRNLGGEVFERLIEP 224
                      ||||||                 | |  |:|  | :| || ||    ::
Db      116 ----------LRPSPPFSKPLFWAGLRELLKPRGKEPDETVHSFAQRRLGPEVASLAMDS 165

Qy      225 FCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKG 284
             | ||: |:  :||::: |  ::: |:|  ||: |      :             |:|
Db      166 LCRGVFAGNSRELSIRSCFPSLFQAEQTHRSILLGLLLGAGQS------------PQPDS 213

Qy      285 QTV----------GSFRKGLRMLPDAISARLGSK------------LKLS----WKLSSI 318
              :           | | || :|| |:   | ||            | |     ||:|
Db      214 SLIRQARAERWSQWSLRGGLEVLPQALHNHLASKGVTVLSGQPVCGLSLQPEGRWKVSL- 272

Qy      319 TKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGA 378
                            |  ||::  |:  :|:   |  :|   :   |  ||      |
Db      273 ----------------GDSSLEADHIISAIPASELSKLLPAEAAPLARILSTIKAVSVAV 316

Qy      379 VTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNR---APKGRVLLL 435
            | : |    :         ::||| | | ::    || :| |  || :    |  || ::
Db      317 VNLQYRGACL--------PVQGFGHLVPSSEDPTVLGIVYDSVAFPEQDGNPPSLRVTVM 368

Qy      436 NYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQD-----------PLVVGVRVWPQ 484
              :||            | ::     |   | : :||:           |    | :
Db      369 --LGG---------YWLQKLKAAGHQLSPELFQQQAQEAAATQLGLKEPPSHCLVHLHKN 417

Qy      485 AIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540
             |||: :||   | :|   :    |     | :|  |||:  |:|   : |  | |
Db      418 CIPQYTIGHWQKLDSAMQFLTAQRLPLTLAGASY-EGVAVNDCIESGRQAAVAVLG 472

RESULT   9
YAM7_SCHPO
ID   YAM7_SCHPO     STANDARD;      PRT;   490 AA.
AC   Q10062;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   HYPOTHETICAL 53.5 KDA PROTEIN C1F5.07C IN CHROMOSOME I.
GN   SPAC1F5.07C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Gentles S., Odell C., Churcher C.M., Barrell B.G., Rajandream M.A.,
RA   Walsh S.V.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) AND TO
CC       YEAST YER014W.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z68136; CAA92235.1; -.
```

```
KW   Hypothetical protein.
SQ   SEQUENCE   490 AA;  53461 MW;  CB679F6A437547AA CRC64;

  Query Match             10.6%;  Score 302.5;  DB 1;  Length 490;
  Best Local Similarity   22.9%;  Pred. No. 2.1e-15;
  Matches  120;  Conservative  103;  Mismatches  213;  Indels   87;  Gaps   20;

Qy     72 LDCVIVGAGISGLCIAQVMSANYPNLMVT--EARDRAGGNITTVE-------RDGYLWEE 122
          :   | | ||:||  |  ::   |   :   |   | || : :|:           |:|:
Db      1 MSIAICGGGIAGLSTAFYLARLIPKCTIDLYEKGPRLGGWLQSVKIPCADSPTGTVLFEQ 60

Qy    123 GPNSFQPSD----PMLTMAVDCGLKDDLVLGDPNAP----RFVLWKGKLRPVPSKLTDLP 174
          || : :|:       | :    |::| |:    |:|    |:: :  :|  :||
Db     61 GPRTLRPAGVAGLANLDLISKLGIEDKLLRISSNSPSAKNRYIYYPDRLNEIPS------ 114

Qy    175 FFDLMSIPGKLRAGFGPIGLRPSP-------------PGHEESVEQFVRRNLGGEVFERL 221
               || | :::   |  ||| |                :|||  |:||  |  | :|:
Db    115 -----SILGSIKSIMQP-ALRPMPLAMMLEPFRKSKRDSTDESVGSFMRRRFGKNVTDRV 168

Qy    222 IEPFCSGVYVGDPSKLSMKAA-FGKVWKLEETGGSIIGGTFKAIKER---SSTPKAPRDP 277
          :    :|:| || : ||| :: || : |:|:  |:|  |  :|:  |   |   || :
Db    169 MSAMINGIYAGDLNDLSMHSSMFGFLAKIEKKYGNITLGLIRALLAREILSPAEKALKAA 228

Qy    278 RLPKP---------KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHL 328
           | :|         |  ::  :|::|:  :  :|:  |     :   |:   |:      |
Db    229 LLAEPKTAELSNSMKSTSMFAFKEGIETITLSIADELKKMPNVKIHLNKPAKT------L 282

Qy    329 TYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAI 388
               : :| :  ::    | :  : |:   :|    :  ::  |     | : |
Db    283 VPHKTQSLVDVNGQAYEYVVFANSSRNLENLISCPKMETPTSSVY----VVNVYY----- 333

Qy    389 RDERLVDGELKGFGQLHPR---TQGVETLGTIYSSSL-FPNRAPKGRVLLLNYIGGA--- 441
          :|  ::   ::||| | |          || :: |    |    |  |::    ||:
Db    334 KDPNVL--PIRGFGLLIPSCTPNNPNHVLGIVFDSEQNNPENGSKVTVMM----GGSAYT 387

Qy    442 KNPEILSKTESQLVEVVDRDLRKML-IKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTA 500
          ||  ::     : |    : |:  | |  |   | :    :    |||: ||| | |::
Db    388 KNTSLIPTNPEEAVNNALKALQHTLKISSK---PTLTNATLQQNCIPQYRVGHQDNLNSL 444

Qy    501 KAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLS 543
          |: :  |    : | |:: :||::| |:   :  | ::   ::
Db    445 KSWIEKNMGGRILLTGSWYNGVSIGDCIMNGHSTARKLASLMN 487

RESULT  10
PPOX_MYCTU
ID   PPOX_MYCTU     STANDARD;      PRT;   426 AA.
AC   O53230;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
GN   HEMY OR RV2677C OR MTV010.01C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Sqares R., Sulston J.E.,
RA   Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
```

```
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
CC   -!- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   -!- CAUTION: SEEMS SHORTER IN THE C-TERMINAL THAN CORRESPONDING HEMY
CC       IN OTHER BACTERIA.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AL021186; CAA16020.1; -.
DR   TubercuList; Rv2677c; -.
KW   Porphyrin biosynthesis; Oxidoreductase; Flavoprotein; FAD;
KW   Heme biosynthesis.
FT   NP_BIND      10     15       FAD (POTENTIAL).
SQ   SEQUENCE   426 AA;  44382 MW;  221952BB3977D3B0 CRC64;

  Query Match              7.8%;  Score 221;  DB 1;  Length 426;
  Best Local Similarity   22.8%;  Pred. No. 2.5e-09;
  Matches  108;  Conservative   74;  Mismatches  210;  Indels   82;  Gaps   17;

Qy     76 IVGAGISGLCIA-QVMSANYPNLMVT--EARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
          :|| |||||  | ::  |   :  :|  |  || || : |    |   : |  :|    |
Db      8 VVGGGISGLTSAYRLRQAVGDDATITLFEPADRLGGVLRTEHIGGQPMDLGAEAFVLRRP 67

Qy    133 -MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGP 191
           |  :  : ||  |  |    |   :  : :| |:| :        :: ||    :  |
Db     68 EMPALLAELGL-SDRQLASTGARPLIYSQQRLHPLPPQ-------TVVGIPSSAGSMAGL 119

Qy    192 I-----------GLRPS--PPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLS 238
          :             ||     | : :|   |    | :|  | ::|  |||| |  : :
Db    120 VDDATLARIDAEAARPFTWQVGSDPAVADLVADRFGDQVVARSVDPLLSGVYAGSAATIG 179

Qy    239 MKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLP 298
          ::||   |    : |              :|   | |    |   |   |:   | ::|
Db    180 LRAAAPSVAAALDRGA-------------TSVTDAVRQALPPGSGGPVFGALDGGYQVLL 226

Qy    299 DAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILR 358
          | :  |     :: |  : : : |: |: |  ||       |: :::: ||:   : ::
Db    227 DGLVRR----SRVHWVRARVVQLER-GWVLRDETGG---RWQADAVILAVPAPRLARLVD 278

Qy    359 PLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLG--- 415
           ::     |           | :: |                 |   |   ||   |
Db    279 GIAPRTHAAARQIVSASSAVVALAVPG----------------GTAFPHCSGVLVAGDES 322

Qy    416 ------TIYSSSLFPNRAPKGRVLLLNYIGGAKNPE-ILSKTESQLVEVVDRDLRKMLIK 468
                |: |         :| | ||    |    |  |: :: ||:     ||  : :
Db    323 PHAKAITLSSRKW----GQRGDVALLRLSFGRFGDEPALTASDDQLLAWAADDL--VTVF 376

Qy    469 PKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
            | ||: | || | :|:||:  || | ::  :| :       | : |:|: |:
Db    377 GVAVDPVDVRVRRWIEAMPQYGPGHADVVAELRAGLPPT----LAVAGSYLDGI 426

RESULT  11
PPOX_MYCLE
ID   PPOX_MYCLE     STANDARD;      PRT;   451 AA.
AC   Q50008;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
```

```
GN   HEMY.
OS   Mycobacterium leprae.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1769;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith D.R., Robison K.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
CC   -!- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U15181; AAA62958.1; -.
KW   Porphyrin biosynthesis; Oxidoreductase; Flavoprotein; FAD;
KW   Heme biosynthesis.
FT   NP_BIND      10     15       FAD (POTENTIAL).
SQ   SEQUENCE   451 AA;  46580 MW;  DF76EE1655CA2056 CRC64;


  Query Match             7.6%;  Score 216.5;  DB 1;  Length 451;
  Best Local Similarity   20.6%;  Pred. No. 5.8e-09;
  Matches  105;  Conservative   79;  Mismatches  193;  Indels  133;  Gaps   18;

Qy         76 IVGAGISGLCIA---QVMSANYPNLMVTEARDRAGGNITTVERDGYLWEEGPNSF---QP 129
              :|| |||||  |   :| : :   : : :  || || : |    |   : |  :|   :|
Db          8 VVGGGISGLTAAYRLRVATGDDVAITLFDPGDRLGGVLRTECVGGQPMDLGAEAFLLRRP 67

Qy        130 SDPMLTMAVDCGLKD-----------------------DLVLGDPNAPRFVLWKGKLRPV 166
                | |    : || :                       | | |                :
Db         68 EVPAL--LAELGLSERQRATTDARPLIYSQQRLHSLPPDTVAG----------------I 109

Qy        167 PSKLTDLPFFDLMSIPGKLR----AGFGPIGLRPS--PPGHEESVEQFVRRNLGGEVFER 220
              ||  |        |: | :     |  |   :||    || : :: : |    | :   |
Db        110 PSSAT--------SVAGLVDDATVARIGAEAVRPLSWEPGSDPAMAELVADRFGEQAVAR 161

Qy        221 LIEPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLP 280
              |::|   ||| |  : : ::|    |    : |              :|  :| |    |
Db        162 LVDPLLGGVYAGSAATIGLRAGAPSVAAALDCGA-------------TSLMEAVRQGLPP 208

Qy        281 KPKGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKG--GYHLTYETPEGVVS 338
                 |   |:   | ::| | :  |     :| |  :::   ::|  |: |  :|
Db        209 VAAGPVFGALDGGYQVLIDELVRR----SRLQWVAATVVGLDRGTCGWTLVDDTG---AC 261

Qy        339 LQSRSIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGEL 398
                :  ::: ||:     :|: ::     | |         | :| |::    :    | |
Db        262 WSADGVILAVPAPRLVRLLQQIAPRTVAAASRIVSASSAVVALSVPRDTTFPQN--SGVL 319

Qy        399 KGFGQ--------LHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGG---------A 441
                 |:        |  |  |::                 |   |:    |         |
Db        320 VASGERLRAKAVTLSSRKWGLQ-----------------GDTQLVRLSFGKFGDQVASTA 362

Qy        442 KNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAK 501
               : |:|:   | |  | |  :          ||: | |: |  |:||:  || | ::  :
Db        363 SDDELLAWAVSDLAAVFDVTV----------DPVDVCVQRWIDAMPQYGPGHADLVAEVR 412

Qy        502 AAMNDNGLEGLFLGGNYVSGVALGRCVEGA 531
              | :       | : |::: |: :  |:  |
Db        413 AGLPPT----LVVAGSHMDGIGVPACISAA 438
```

```
RESULT  12
PPOX_YEAST
ID   PPOX_YEAST     STANDARD;      PRT;   539 AA.
AC   P40012;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
GN   HEM14 OR YER014W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96224138; PubMed=8621563;
RA   Camadro J.-M., Labbe P.;
RT   "Cloning and characterization of the yeast HEM14 gene coding for
RT   protoporphyrinogen oxidase, the molecular target of diphenyl
RT   ether-type herbicides.";
RL   J. Biol. Chem. 271:9120-9128(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA   Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA   Chung E., Duncan M., Guzman E., Hartzell G., Hunicke-Smith S.,
RA   Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA   Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA   Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA   Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SUBCELLULAR LOCATION.
RC   STRAIN=ATCC 25657 / D273-10B;
RX   MEDLINE=95096046; PubMed=7798202;
RA   Camadro J.-M., Thome F., Brouillet N., Labbe P.;
RT   "Purification and properties of protoporphyrinogen oxidase from the
RT   yeast Saccharomyces cerevisiae. Mitochondrial location and evidence
RT   for a precursor form of the protein.";
RL   J. Biol. Chem. 269:32085-32091(1994).
CC   -!- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC       IX TO FORM PROTOPORPHYRIN IX. OPTIMUM ACTIVITY IS OBTAINED AT PH
CC       7.2.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O(2).
CC   -!- COFACTOR: FAD; CONTAINS 1 MOLE OF FAD PER MOLE OF ENZYME.
CC   -!- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE
CC       WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE.
CC   -!- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z71381; CAA95981.1; -.
DR   EMBL; U18778; AAB64547.1; -.
DR   SGD; S0000816; HEM14.
KW   Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW   Flavoprotein; FAD; Mitochondrion.
FT   NP_BIND      18     23       FAD (POTENTIAL).
FT   MUTAGEN     422    422       L->P: IN HEM14-1; LOSS OF ACTIVITY.
FT   MUTAGEN     424    424       K->E: IN HEM14-1; LOSS OF ACTIVITY.
SQ   SEQUENCE   539 AA;  59703 MW;  5ECCBD1C033BA5B1 CRC64;
```

```
  Query Match             6.6%;  Score 189;  DB 1;  Length 539;
  Best Local Similarity   20.3%;  Pred. No. 9e-07;
  Matches  111;  Conservative   95;  Mismatches  202;  Indels  138;  Gaps   24;

Qy      68 PAAELDCVIVGAGISGLCIAQVMSANYPNLMVT--EARDRAGG-----NITTVERDGYLW 120
           | |::   :|| |:||||    :|   |:: :|  |:::| ||     |   :  :  :
Db      10 PRAKV--AVVGGGVSGLCFTYFLSKLRPDVEITLFESQNRTGGWIYSCNTRDMSGNPIML 67

Qy     121 EEGPNSFQ-PSD--------------PMLTMAVDCG-LKDDLVLGDPNAPRFVLWKGKLR 164
           |:|| : :  ||                :  ::| | : |   | ||:         ||
Db      68 EKGPRTLRGVSDGTVLIMDTLKDLGKEAVIQSIDKGCIADKKFLLDPS--------DKLV 119

Qy     165 PVPSKLTDLPFFDLMSI-PGKLRAGFGPIGLRPSP-PGHEESVEQFVRRNLGGE-VFERL 221
            ||: ::    | |  :  | :    |    : || || :||||    |  |   :   :
Db     120 QVPNSISTTVKFLLNPLGKGLITGMMGEWFRKKSPHPGQDESVESICDRRFGNNYISNNM 179

Qy     222 IEPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPR------ 275
           |     |:|  | | || |  | |::  |   ||        :: :| : |
Db     180 ISALLRGIYGDDVSLLSAKRTFKKIYYNELKHGSNTQAMIDNMRGKSRSKKTENLHQSLT 239

Qy     276 ------DPRLPKPKGQTV---GSFRK--------GLRMLPDAISARLGS----KLKLSWK 314
                      | : : :    : :|        ||   |  :   |      |:
Db     240 GCLNDYSNAFGKDRSKLLDLSNTLKKYPMLGLAGGLETFPKIVRNALNEFKNVKIVTGNP 299

Qy     315 LSSI-----------TKSEKGGYHLTYE------TPEGVVSL--QSRSIVMTVPSYVASN 355
           :: |            |:: |  : |::      ||  :  |  : :: :  :   : ||
Db     300 VTQIMKRPANETTIGLKAKSGDQYETFDHLRLTITPPKIAKLLPKDQNSLSKLLDEIQSN 359

Qy     356 ILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRT--QGVET 413
            :          | |:| |               :: ::| :|:||| | |::     :
Db     360 TI---------ILVNYYLP---------------NKDVIDADLQGFGYLVPKSNKNPGKL 395

Qy     414 LGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSK-------------TESQLVEVVDR 460
           || |: | :  |  |     | : :  : ||  |:|              |  :  |  |
Db     396 LGVIFDSVIERNFKP-----LFDKL--STNPNALNKYTKVTAMIGGCMLNEHGVPVVPSR 448

Qy     461 DLRKMLIKPKAQDPLVVGVR----------VWPQAIPQFLVGHLDTLSTAKAAMNDNGLE 510
           ::    :|    : | :  :          :  : :|:| ||:      |:  : ::  :
Db     449 EVTINAVKDALNNHLGISNKDLEAGQWEFTIADRCLPRFHVGYDAWQERAERKLQESYGQ 508

Qy     511 GLFLGG 516
            : :||
Db     509 TVSVGG 514

RESULT  13
AOFH_MYCTU
ID   AOFH_MYCTU     STANDARD;      PRT;   448 AA.
AC   O53320;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   PUTATIVE FLAVIN-CONTAINING MONOAMINE OXIDASE RV3170 (EC 1.4.3.-).
GN   RV3170 OR MTV014.14.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Sqares R., Sulston J.E.,
RA   Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
```

```
CC   -!- COFACTOR: FAD (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AL021646; CAA16635.1; -.
DR   TubercuList; Rv3170; -.
DR   InterPro; IPR002937; -.
DR   Pfam; PF01593; Amino_oxidase; 1.
KW   Hypothetical protein; Oxidoreductase; Flavoprotein; FAD.
FT   NP_BIND       8     63       FAD (ADP PART) (POTENTIAL).
SQ   SEQUENCE   448 AA;  48439 MW;  93FA2C74B3F13D7A CRC64;

  Query Match             5.0%;  Score 141;  DB 1;  Length 448;
  Best Local Similarity   21.6%;  Pred. No. 0.0029;
  Matches   65;  Conservative   59;  Mismatches  119;  Indels   58;  Gaps   13;

Qy      68 PAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLWEEGPNSF 127
           |   :| |:|||| :||  |: ::     ::| | ||| ||   |    |   : | :
Db       4 PPWTVDVVVVGAGFAGLAAARELTRQGHEVLVFEGRDRVGGRSLTGRVAGVPADMGGSFI 63

Qy     128 QPS-DPMLTMAVDCGLKDDLVLGDPNAP------RFVLWKGKLRPVPSKLTDLPFFDLMS 180
            |: | :| :| :        || |  |        : |:|  |     :  |    |:
Db      64 GPTQDAVLALATE--------LGIPTTPTHRDGRNVIQWRGSARSYRGTIPKLSLTGLID 115

Qy     181 IPGKLRAGFGPI--GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLS 238
           | |:||  |  |  |:  : |       :    :||   : ||:            |  :
Db     116 I-GRLRWQFERIARGVPVAAPWDARRARELDDVSLG--EWLRLVR--------ATSSSRN 164

Qy     239 MKAAFGKV-WKLEETGGSII---------GGTFKAIKERSSTPKAPRDPRLPKPKGQTVG 288
           : |   :| |  |    |::         ||  : :  ::      :  |:|
Db     165 LMAIMTRVTWGCEPDDVSMLHAARYVRAAGGLDRLLDVKNGA----QQDRVP-------- 212

Qy     289 SFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTV 348
               | : :  | :|:||::: |:   ::: : :: |  :|  : :|    ::  ::: :
Db     213 ---GGTQQIAQAAAAQLGARVLLN---AAVRRIDRHGAGVTVTSDQG--QAEAGFVIVAI 264

Qy     349 P 349
           |
Db     265 P 265

RESULT  14
CRTI_NARPS
ID   CRTI_NARPS     STANDARD;      PRT;   570 AA.
AC   Q40406;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PHYTOENE DEHYDROGENASE PRECURSOR (EC 1.3.-.-) (PHYTOENE DESATURASE).
GN   PDS1 OR PDS.
OS   Narcissus pseudonarcissus (Daffodil).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae; Narcissus.
OX   NCBI_TaxID=39639;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Paracorolla;
RA   Al-Babili S., Beyer P.;
RT   "A cDNA encoding phytoene desaturase from daffodil.";
RL   (In) Plant Gene Register PGR95-131.
RN   [2]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   TISSUE=Paracorolla;
RX   MEDLINE=96237452; PubMed=8653112;
RA   Al-Babili S., von Lintig J., Haubruck H., Beyer P.;
```

```
RT   "A novel, soluble form of phytoene desaturase from Narcissus
RT   pseudonarcissus chromoplasts is Hsp70-complexed and competent for
RT   flavinylation, membrane association and enzymatic activation.";
RL   Plant J. 9:601-612(1996).
RN   [3]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE=97433278; PubMed=9288918;
RA   Bonk M., Hoffmann B., von Lintig J., Schledz M., Al-Babili S.,
RA   Hobeika E., Kleinig H., Beyer P.;
RT   "Chloroplast import of four carotenoid biosynthetic enzymes in vitro
RT   reveals differential fates prior to membrane binding and oligomeric
RT   assembly.";
RL   Eur. J. Biochem. 247:942-950(1997).
CC   -!- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE
CC       INTERMEDIARY OF PHYTOFLUENE BY THE SYMMETRICAL INTRODUCTION OF TWO
CC       DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.
CC   -!- COFACTOR: FAD (PROBABLE).
CC   -!- PATHWAY: CAROTENOID BIOSYNTHESIS.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST (CHROMOPLAST). EXISTS AS AN
CC       INACTIVE SOLUBLE FORM AND AN ACTIVE MEMBRANE-BOUND FORM.
CC   -!- TISSUE SPECIFICITY: EXPRESSED MORE STRONGLY IN FLOWERS THAN IN
CC       LEAVES.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSION INCREASES DURING FLOWER
CC       DEVELOPMENT.
CC   -!- SIMILARITY: TO OTHER PLANTS OR CYANOBACTERIAL PHYTOENE
CC       DESATURASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X78815; CAA55392.1; -.
DR   Mendel; 13184; Narps;Pds1;13184.
DR   InterPro; IPR000205; -.
DR   InterPro; IPR002937; -.
DR   Pfam; PF01593; Amino_oxidase; 1.
KW   Carotenoid biosynthesis; Oxidoreductase; FAD; Chloroplast; Membrane;
KW   Transit peptide.
FT   TRANSIT       1     91       CHLOROPLAST (POTENTIAL).
FT   CHAIN        92    570       PHYTOENE DEHYDROGENASE.
FT   NP_BIND     104    120       FAD (ADP PART) (BY SIMILARITY).
SQ   SEQUENCE   570 AA;  63790 MW;  654F569F5B83BE77 CRC64;

  Query Match             4.9%;  Score 139;  DB 1;  Length 570;
  Best Local Similarity   21.1%;  Pred. No. 0.0058;
  Matches  111;  Conservative   92;  Mismatches  178;  Indels  146;  Gaps   24;

Qy      68 PAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITT-VERDGYLWEEGPNS 126
           |   |: |:||||::||  |: ::      :: |:||  || |    ::||  :| | :
Db      94 PEKGLEVVVVGAGLAGLSTAKYLADAGHKPILLESRDVLGGKIAAWKDKDGDWYETGLHI 153

Qy     127 FQPSDP-MLTMAVDCGLKD-------DLVLGDPNAP----RFVLWKGKLRPVPSKLTDLP 174
           |  : | :  :  : |: |        ::   || |    ||   :    |:      |
Db     154 FFGAYPNVQNLFGELGINDRLQWKEHSMIFAMPNKPGEFSRFDFPEVLPAPLNGIWAILR 213

Qy     175 FFDLMSIPGKLRAGFGPIGLRPSPPGHEESVE--------QFVRR-----NLGGEVFERL 221
             :::: | |:|     ||| |:  | :  ||        :::||      :  |||  :
Db     214 NNEMLTWPEKVRFA---IGLLPAMVGGQAYVEAQDGLTVTEWMRRQGVPDRVNDEVFIAM 270

Qy     222 IEPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPK 281
                  :   :| :|||:               |:    : ::|:  :  |  |   |:
Db     271 ----SKALNFINPDELSMQC--------------ILIALNRFLQEKHGSKMAFLDGNPPE 312

Qy     282 PKGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVS--L 339
                     :    : | | : || : :|: :|  |  :           |:| |   :
Db     313 ----------RLCMPIVDHIQS-LGGRAQLNSRLQKIELN-----------PDGTVKHFV 350

Qy     340 QSRSIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQE--AIRDERLVDGE 397
                ::|  :||         |||          ||  : :  |||   |   : :| :
Db     351 LGNGNIITGDAYV---------VAA----------PVDILKLLLPQEWREIPYFQKLD-K 390

Qy     398 LKGFGQLHPRTQGVETLGTIYSSSLF---------------------PNRAPKGRVLLLN 436
           | |   ::        |   |   ||                     |||:    :| |
Db     391 LVGVPVINVHIWFDRKLKNTYDHLLFTRSPLLSVYADMSVTCKEYYDPNRS----MLELV 446

Qy     437 YIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDT 496
           :   |   | :|:::|:::|   ::| |:     | |          |:  : |  |:
Db     447 F---APAEEWISRSDSEIIERTMKELAKLFPDEIAAD----------QSKAKILKYHV-- 491

Qy     497 LSTAKAA------------MNDNGLEGLFLGGNYVSGVALGRCVEGA 531
           : | ::             :  : :|| :| |:| :   |   :|||
Db     492 VKTPRSVYKTIPDCEPCRPLQRSPIEGFYLAGDYTNQKYLA-SMEGA 537


RESULT  15
AOF_ONCMY
ID   AOF_ONCMY      STANDARD;      PRT;   522 AA.
AC   P49253;
DT   01-FEB-1996 (Rel. 33, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   AMINE OXIDASE [FLAVIN-CONTAINING] (EC 1.4.3.4) (MONOAMINE OXIDASE)
DE   (MAO).
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=95107295; PubMed=7808446;
RA   Chen K., Wu H.-F., Grimsby J., Shih J.C.;
RT   "Cloning of a novel monoamine oxidase cDNA from trout liver.";
RL   Mol. Pharmacol. 46:1226-1233(1994).
RN   [2]
RP   IDENTIFICATION OF PROBABLE FRAMESHIFT.
RA   Bairoch A.;
RL   Unpublished observations (JUL-1998).
CC   -!- FUNCTION: OXIDIZES BOTH 5-HYDROXYTRYPTAMINE (5-HT) AND BETA-
CC       PHENYLETHYLAMINE (PEA).
CC   -!- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
CC       H(2)O(2).
CC   -!- COFACTOR: FAD.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL OUTER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC   -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN IN THAT, DUE TO A
CC       PROBABLE FRAMESHIFT, IT ENDS IN POSITION 499 AND LACKS THE C-
CC       TERMINAL REGION FOUND IN OTHER MEMBERS OF THIS FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L37878; AAA64302.1; ALT_FRAME.
DR   InterPro; IPR001613; -.
DR   InterPro; IPR002937; -.
DR   Pfam; PF01593; Amino_oxidase; 1.
DR   PRINTS; PR00757; AMINEOXDASEF.
KW   Oxidoreductase; Flavoprotein; FAD; Transmembrane; Mitochondrion.
FT   NP_BIND       7     63       FAD (ADP PART) (POTENTIAL).
FT   BINDING     399    399       FAD (BY SIMILARITY).
FT   TRANSMEM    493    513       POTENTIAL.
SQ   SEQUENCE   522 AA;  58937 MW;  F82D6D15364D646D CRC64;

  Query Match             4.7%;  Score 134;  DB 1;  Length 522;
```

```
  Best Local Similarity   22.0%;  Pred. No. 0.012;
  Matches   67;  Conservative   48;  Mismatches  119;  Indels   70;  Gaps   12;

Qy     73 DCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLWEE------GPNS 126
          | :::| |||||  |:::     : :| ||||| ||   ||: :   : :      ||
Db      8 DVIVIGGGISGLSAAKLLKEKGLSPVVLEARDRVGGRTFTVQNEQTKYVDLGGAYVGPT- 66

Qy    127 FQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDL--PF--------- 175
              : :| :| :||:|   |  :     :|  |||  |       :  ||
Db     67 ---QNRILRLAKECGVKTIKVNEEERLVHYV--KGKSYPFKGSFPPMWNPFALMDYNNLW 121

Qy    176 --FDLM--SIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEV-----FERLIEPFC 226
             | |   || :           |    | |  ::   : |  ::       |    |
Db    122 RKMDEMGSEIPRE----------APWKAPHAEEWDKMTMKQLFDKICWTSSARRFATLFV 171

Qy    227 SGVYVGDPSKLSMKAAFGKVWKLEETGGSI-IGGTFKAIKERSSTPKAPRDPRLPKPKGQ 285
          :     :| ::|   |   :|  ::: ||:: |  |    :||
Db    172 NVNVTSEPHEVS---ALWFLWYVKQCGGTMRIFSTTNGGQER------------------ 210

Qy    286 TVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIV 345
              |  |   : : ::  || ::|:    | : | :: |  :  ||           ||
Db    211 ---KFLGGSSQISECMAKELGERVKME---SPVYKIDQTGDMVEVETLNKETYKAKYVIV 264

Qy    346 MTVP 349
           | |
Db    265 ATPP 268


Search completed: July  3, 2001, 10:43:19
Job time: 612 sec
```

```
OM of: US-09-508-418-2 to: EST:*   out_format : pfs

Date: Jul 3, 2001 11:06 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09508418/runat_03072001_100016_24427/app_query.fasta_1.613
 -DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09508418_@CGN1_1_3324 -NCPU=6
 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-508-418-2
Query length: 548
Database: EST:*
Database sequences: 10228115
Database length: 431459454
Search time (sec): 1358.990000


score_list:
Sequence              Strd Orig    ZScore   EScore Len  ! Documentation  ..
gb_est73:BE343078      +  891.50 1639.28  3.4e-82   593  ! BE343078 EST395922 potato stolo
gb_est50:AW695211      +  838.00 1538.73  1.4e-76   640  ! AW695211 NF092G06ST1F1051 Devel
gb_est99:BG320666      +  822.50 1507.46  7.6e-75   780  ! BG320666 Zm04_06e01_R Zm04_AAFC
gb_est99:BG352959      +  797.00 1463.39  2.2e-72   590  ! BG352959 sab93f10.y1 Gm-c1040 G
gb_est74:BE455801      +  778.00 1428.22  2.0e-70   580  ! BE455801 HVSMEg0015D05f Hordeum
gb_est39:AV538014      -  739.00 1355.28  2.3e-66   596  ! AV538014 AV538014 Arabidopsis t
gb_est90:BF624024      +  685.00 1256.64  7.1e-61   510  ! BF624024 HVSMEa0008L16f Hordeum
gb_est42:AW119581      +  674.00 1236.44  9.4e-60   498  ! AW119581 sd48d01.y1 Gm-c1016 Gl
gb_est69:BE021686      +  629.00 1153.11  4.1e-55   480  ! BE021686 sm61b05.y1 Gm-c1028 Gl
gb_est99:BG355389      +  621.00 1138.24  2.8e-54   479  ! BG355389 947043H05.y2 947 - 2 w
gb_est39:AV540305      -  605.00 1107.47  1.4e-52   519  ! AV540305 AV540305 Arabidopsis t
gb_est24:AI748531      +  593.00 1086.68  2.1e-51   457  ! AI748531 sb54d01.y1 Gm-c1016 Gl
gb_est97:BG133906      +  589.00 1079.70  5.1e-51   440  ! BG133906 EST466798 tomato crown
gb_est73:BE326089      +  552.50 1008.16  4.9e-47   590  ! BE326089 NF084G10ST1F1081 Devel
gb_est51:AW755783      +  543.00  993.75  3.1e-46   451  ! AW755783 sl09c11.y1 Gm-c1036 Gl
gb_est70:BE129089      +  534.00  974.95  3.5e-45   533  ! BE129089 894020B03.y1 C. reinha
gb_est44:AW202400      +  525.00  961.28  2.0e-44   414  ! AW202400 sf14e09.y1 Gm-c1027 Gl
gb_est42:AW091632      +  492.00  902.35  3.8e-41   337  ! AW091632 EST284908 tomato mixed
gb_est39:AV551750      +  491.00  893.98  1.1e-40   574  ! AV551750 AV551750 Arabidopsis t
gb_est27:AI997032      -  460.50  840.87  1.0e-37   425  ! AI997032 701551755 A. thaliana,
gb_est72:BE229685      +  431.00  783.24  1.7e-34   530  ! BE229685 98SN0073 Rice Immature
gb_est101:BG462577     -  419.00  761.30  2.8e-33   513  ! BG462577 947058C08.x1 947 - 2 w
gb_est39:AV550160      +  416.00  757.42  4.5e-33   446  ! AV550160 AV550160 Arabidopsis t
gb_est11:AA712377      +  415.00  756.47  5.1e-33   414  ! AA712377 32105 Lambda-PRL2 Arab
gb_est39:AV540703      -  408.00  744.26  2.5e-32   387  ! AV540703 AV540703 Arabidopsis t
gb_est27:AI941783      -  398.00  723.11  3.7e-31   476  ! AI941783 618035B09.x1 618 - Inh
gb_est47:AW453355      +  394.00  714.99  1.0e-30   503  ! AW453355 660040G08.y1 660 - Mix
gb_est46:AW428647      -  388.00  702.08  5.5e-30   580  ! AW428647 660040G08.x2 660 - Mix
gb_est76:BE601207      +  365.00  659.67  1.3e-27   561  ! BE601207 PI1_90_C06.g1_A002 Pat
gb_est31:AU175016      +  359.00  652.23  3.3e-27   413  ! AU175016 AU175016 Rice cDNA fro
gb_est101:BG462578     +  357.50  653.38  2.8e-27   299  ! BG462578 947058C08.y1 947 - 2 w
gb_est53:AW929479      +  351.00  639.50  1.7e-26   346  ! AW929479 EST338267 tomato flowe
gb_est31:AU175015      +  276.00  495.78  1.7e-18   483  ! AU175015 AU175015 Rice cDNA fro
gb_est51:AW781134      +  272.00  489.45  3.8e-18   441  ! AW781134 sl89e10.y1 Gm-c1037 Gl
gb_est45:AW313299      -  257.00  460.90  1.5e-16   464  ! AW313299 660023B03.x1 660 - Mix
gb_est49:AW615911      +  253.50  451.51  5.0e-16   587  ! AW615911 EST325409 tomato flowe
gb_est100:BG453937     +  251.50  446.34  9.6e-16   661  ! BG453937 NF096C05LF1F1035 Devel
gb_est109:T43573       +  246.00  440.26  2.1e-15   470  ! T43573 6836 Lambda-PRL2 Arabido
gb_gss19:AQ954144      -  220.00  393.08  8.9e-13   425  ! AQ954144 nbeb0054K15r CUGI Rice
gb_est90:BF598346      +  215.50  381.03  4.2e-12   574  ! BF598346 sv16e03.y1 Gm-c1056 Gl
gb_est50:AW695910      +  215.50  380.66  4.4e-12   592  ! AW695910 NF099H07ST1F1063 Devel
gb_gss17:AQ858291      -  214.00  381.65  3.9e-12   434  ! AQ858291 nbeb0012I22f CUGI R
gb_est51:AW755660      +  213.50  376.97  7.0e-12   590  ! AW755660 sl06e04.y1 Gm-c1036
em_estpl5:AW726965     +  207.50  361.03  5.4e-11   872  ! Aw726965 GA__Ea0023F07 Gossy
gb_est100:BG444110     +  207.50  361.03  5.4e-11   872  ! BG444110 GA__Ea0023F07f Goss

seq_name: gb_est73:BE343078

seq_documentation_block:
LOCUS       BE343078      593 bp    mRNA            EST       17-JUL-2000
DEFINITION  EST395922 potato stolon, Cornell University Solanum tuberosum cDNA
            clone cSTA22L8, mRNA sequence.
ACCESSION   BE343078
VERSION     BE343078.1  GI:9252610
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 593)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Bachem,C., Horvath,B., Visser,R.,
            Holt,I.E., Liang,F., Hansen,T.S., Utterback,T., Bowman,C.L., Doan
            ,B., Bougri,O., Buell,C.R., Ronning,C.M., Tanksley,S.D. and Baker
            ,B.
  TITLE     Generation of ESTs from potato swelling stolons
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Research Genetics, Libraries Division
            Tel: 1-800-711-6195
            Email: cdna@resgen.com
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1..593
                     /organism="Solanum tuberosum"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
                     /clone="cSTA22L8"
                     /clone_lib="potato stolon, Cornell University"
                     /tissue_type="axillary buds of stem explants, swelling
                     stolons"
                     /dev_stage="1 to 3 days"
                     /lab_host="SOLR"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoRl; Site_2:
                     Xhol; RNA was supplied by Christian Bachem & Beatrix
                     Horvath(Laboratory of Plant Breeding, Dept. of Plant
                     Sciences,  Wageningen University, The Netherlands). Total
                     RNA was isolated from developing axillary buds of potato
                     nodal stem cuttings  cultured on medium for the
                     introduction of tuber formation as  described in Bachem et
                     al. (Plant Journal 1996). Tissue samples were  taken of
                     stages corresponding to growing stolons and the early
                     stages  of tuber formation."
BASE COUNT      129 a    129 c    158 g    177 t
ORIGIN


alignment_scores:
          Quality:  891.50              Length:    191
            Ratio:  4.980                 Gaps:      1
 Percent Similarity:  93.717   Percent Identity: 88.482

alignment_block:
 US-09-508-418-2 x BE343078    ..

 Align seg 1/1  to: BE343078  from: 1  to: 593

      17 SerSerSerSerPro...LeuAlaPheLeuAsnArgThrSerPheIlePr 32
         |||||||||||||||   :::   |||||||||||||||:::||||||||
      21 TCCTCCTCATCGCCGTCATTTTTATTTTTAAATCGTACGAATTTCATTCC 70

      32 oPheSerSerIleSerLysArgAsnSerValAsnCysAsnGlyTrpArgT 49
         |:::   |||   |||||||||||||||||||||||||||||||||||||
      71 TTACTTTTCCACCTCCAAGCGCAATAGTGTCAATTGCAATGGCTGGAGAA 120
```

```
   49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
      |||||||||||||||||||||||||||||||||||   |||   ||||||
  121 CACGATGTTCCGTTGCCAAGGATTATACAGTTCCTCCCTCGGAAGTCGAC 170

   66 GlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSe 82
      |||:::         |||||||||||||||:::|||||||||||||||||
  171 GGTAATCAGTTCCCGGAGCTGGATTGTGTGGTAGTTGGAGCAGGAATTAG 220

   82 rGlyLeuCysIleAlaGlnValMetSerAlaAsnTyrProAsnLeuMetV 99
      ||||||||||||||||:::|||:::|||   |||||||||||||||||||
  221 TGGACTCTGCATTGCTAAGGTGATTTCGGATAATTATCCCAATTTGATGG 270

   99 alThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArg 115
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  271 TGACGGAGGCGAGGGATCGTGCCGGTGGAAACATAACGACGGTGGAAAGA 320

  116 AspGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPr 132
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  321 GATGGATACTTATGGGAAGAAGGTCCTAACAGTTTCCAGCCTTCGGATCC 370

  132 oMetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuG 149
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  371 TATGTTGACAATGGCTGTAGATTGTGGATTGAAGGATGATTTGGTGTTGG 420

  149 lyAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgPro 165
      ||||||||:::||||||||||||||||||||||||   ||||||||||||
  421 GAGATCCTGATGCGCCTCGCTTTGTCTTGTGGAAGGATAAACTAAGGCCT 470

  166 ValProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePr 182
      ||||||:::|||||||||||||||||||||||||||||||||||||||||
  471 GTTCCCGGCAAGCTCACTGATCTTCCCTTCTTTGATTTGATGAGTATCCC 520

  182 oGlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProP 199
      |||||||||||||||||||||||||   ||||||||||||||||||||||
  521 TGGCAAGCTCAGAGCTGGTTTTGGTGCCATTGGCCTTCGCCCTTCACCTC 570

  199 roGlyHisGluGluSerValGlu 206
      |||||:::|||||||||||||||
  571 CAGGTTATGAGGAATCAGTTGAG 593

seq_name: gb_est50:AW695211

seq_documentation_block:
LOCUS       AW695211        640 bp    mRNA            EST       15-JUN-2000
DEFINITION  NF092G06ST1F1051 Developing stem Medicago truncatula cDNA clone
            NF092G06ST 5', mRNA sequence.
ACCESSION   AW695211
VERSION     AW695211.1  GI:7569973
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE   1  (bases 1 to 640)
  AUTHORS   He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
            ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
            ,R.A.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula stem library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Dixon RA
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7302
            Fax: 580 221 7380
            Email: radixon@noble.org
            Insert Length: 640   Std Error: 0.00
            Plate: 092  row: G  column: 06
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF092G06ST"
                     /clone_lib="Developing stem"
                     /tissue_type="stem"
                     /dev_stage="Pooled developmental"
                     /note="Vector: Lambda Zap; Contains a mixture of
                     internodal stem segments"
BASE COUNT      174 a    126 c    146 g    194 t
ORIGIN

alignment_scores:
          Quality:  838.00            Length:    214
            Ratio:  4.169               Gaps:      1
 Percent Similarity:  93.925   Percent Identity: 76.636

alignment_block:
 US-09-508-418-2 x AW695211   ..

 Align seg 1/1  to: AW695211  from: 1  to: 640

  300 AlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSe 316
      |||||||||||||||||||||:::|||:::||||||||||||||||||
    1 GCAATTTCTGCTAGGTTAGGTAACAAGGTGAAGTTATCTTGGAAGCTTTT 50

  316 rSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrP 333
       ||||||:::|||   ::::::|||   |||   ||||||||||||||||
   51 AAGTATTAGTAAACTGGATAGTGGAGAGTACAGTTTGACGTATGAAACCC 100

  333 roGluGlyValValSerLeuGlnSerArgSerIleValMetThrValPro 349
      ||||||||||||||||||||||||||::::::::::::||||||:::|||
  101 CAGAAGGAGTTGTTTCTTTGCAGAGCAAGACTGTTATCATGACCATTCCT 150

  350 SerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAs 366
      |||:::|||||||||   :::|||||||||||||||   |||||||||||
  151 TCCCATGTTGCTAGTCCATTGCTGCGACCACTGTCTCCTGCTGCTGCAGA 200

  366 pAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerT 383
      ||||||||||:::||||||||||||||||||:::||||||:::|||||||
  201 TGCACTTTCAAAGTTTTATTATCCTCCAGTTGCCGCAGTTTCCATTTCCT 250

  383 yrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLys 399
      |||||:::||||||||||||:::|||   |||:::|||||||||||||||
  251 ATCCAAAAGAAGCCATTAGATCAGAATGCTTGATAGATGGCGAGTTGAAG 300

  400 GlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyTh 416
      ||||||||||||||||||||||||:::|||||||||:::|||||||||||
  301 GGATTTGGTCAATTGCATCCACGGAGCCAAGGAGTGCAAACATTAGGAAC 350

  416 rIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgValL 433
      |||||||||||||||||||||||||||||||||||||   ||||||||||
  351 TATATACAGCTCATCCCTTTTCCCTAACCGAGCACCACCTGGAAGGGTTC 400

  433 euLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLys 449
      ||||||||||||||||||||||||||   |||      ||||||||||||
  401 TACTCTTGAATTACATTGGAGGGGCTACCAATTCTGGGATTTTATCAAAG 450

  450 ThrGluSerGlnLeuValGluValValAspArgAspLeuArgLysMetLe 466
      ||||||::::::|||||||||:::|||||||||||||||||||||:::||
  451 ACGGAGGGTGAACTTGTTGAAGCAGTTGATCGAGATTTGAGAAAAATCCT 500

  466 uIleLysProLysAlaGlnAspProLeuValValGlyValArgValTrpP 483
      ||||:::|||:::||||||||||||:::|||:::||||||||||||||||
  501 TATAAACCCAAATGCTCAGGATCCAATTGTTTTGGGGGTTAGAGTGTGGC 550

  483 roGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThr 499
      ||||||||||||||||||||||||||||||||||||||   |||::::::
  551 CTCAAGCTATTCCACAGTTCTTAGTTGGACATCTTGATCTTCTAGATGCT 600

  500 AlaLysAlaAlaMet.AsnAspAsnGlyLeuGluGlyLeu 512
      ||||||::::::::: :::::::::|||:::|||||||||
  601 GCTAAAGTTTCTCTAAAAAAATACTGGATTTGAGGGGCTT 640
```

```
seq_name: gb_est99:BG320666

seq_documentation_block:
LOCUS       BG320666       780 bp    mRNA            EST       27-FEB-2001
DEFINITION  Zm04_06e01_R Zm04_AAFC_ECORC_cold_stressed_maize_seedlings Zea mays
            cDNA clone Zm04_06e01, mRNA sequence.
ACCESSION   BG320666
VERSION     BG320666.1  GI:13150344
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 780)
  AUTHORS   Singh,J.A., Wakui,K., Couroux,P., De Moors,A., Harris,L.J., Hattori
            ,J.I., Ouellet,T., Robert,L.S., Sprott,D. and Tinker,N.A.
  TITLE     Expressed Sequence Tags from Cold-Stressed Maize Seedlings
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Singh,J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            960 Carling Avenue, Bldg. 20, Ottawa, Ontario, K1A 0C6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
            Email: singhja@em.agr.ca.
FEATURES             Location/Qualifiers
     source          1. .780
                     /organism="Zea mays"
                     /cultivar="CO328"
                     /db_xref="taxon:4577"
                     /clone="Zm04_06e01"
                     /clone_lib="Zm04_AAFC_ECORC_cold_stressed_maize_seedlings"
                     /tissue_type="Leaf, crown"
                     /note="Vector: Bluescript SK-/XhoI-EcoRI; Site_1: Eco RI;
                     Site_2: Xho I; Lower temperature 5o C / hour from 22 to
                     12oC; bring to 5o in 1 hour from 12oC. Leave at 5oC 2 days
                     , photoperiod 16 hours. Light intensity was 125 uE-1.
                     Library prepared by in vivo mass excision from amplified
                     library."
BASE COUNT      185 a    198 c    186 g    175 t     36 others
ORIGIN


alignment_scores:
           Quality:  822.50             Length:    258
             Ratio:  3.880                Gaps:      2
 Percent Similarity:  82.171   Percent Identity: 67.054

alignment_block:
 US-09-508-418-2 x BG320666   ..

 Align seg 1/1  to: BG320666  from: 1  to: 780

    123 GlyProAsnSerPheGlnProSerAspProMetLeuThrMetAlaValAs 139
        ||||||||||||   ||||||:::||||||:::|||||||||||||||||
      1 GGTCCCAACAGCYTCCAGCCCYCCGACCCCGTTCTCACCATGGCCGTTGA 50

    139 pCysGlyLeuLysAspAspLeuValLeuGlyAspProAsnAlaProArgP 156
        |   |||||||||||||||||||||   ||||||||||||||||||||||
     51 CAGCGGACTGAAGGATGACTTGGTTKTTGGGGACCCAAACGCGCCGCGTT 100

    156 heValLeuTrpLysGlyLysLeuArgProValProSerLysLeuThrAsp 172
        ||||||||   :::||||||      |||||||||:::|||   :::|||
    101 TCGTGCTGTKGGAGGGGAAGCYGARGCCCGTWCCAACCAARCCCGCCGAC 150

    173 LeuProPhePheAspLeuMetSerIleProGlyLysLeuArgAlaGlyPh 189
        |||||||||||||||||||||:::||||||||||||||||||||||||::
    151 CTCCCGTTCTTCGATCTCATGARCATCCCAGGGAAGCTCAGGGCCGGKCT 200

    189 eGlyProIleGlyLeuArgProSerProProGlyHisGluGluSerValG 206
        :|||         :::||||||   |||||||||::::::|||:::   |
    201 AGGCGCGCYTTGCATCCGCCCGCCTCCYCCAGGCCGCCAAGAGTYARTTG 250

    206 luGlnPheValArgArgAsnLeuGlyGlyGluValPheGluArgLeuIle 222
        ||:::|||||||||||||||   |||:::|||   ||||||   |||
    251 AGGAGTTCGTTCGCCGCAACCYTGGTGCTGAGGKCTTCGAGCSCCTCAAT 300

    223 GluProPheCysSerGlyValTyrValGlyAspProSerLysLeuSerMe 239
           ||||||||||||   ||||||   |||||||||||||||   |||||
    301 KAGCCTTTCTGCTCARGTGTCTATTCTGGTGATCCTTCTAARCYCAGCAT 350

    239 tLysAlaAlaPheGlyLysValTrpLysLeuGluGluThrGlyGlySerI 256
        ||||||||||||||||||||||   :::||||||||||||   |||||||
    351 GAAGGCTGCATTTGGGAAGGTTTKGCGGTTGGAAGAAACTKGAGGTAGTA 400

    256 leIleGlyGlyThrPheLysAlaIleLysGluArgSerSerThrProLys 272
        |||||   |||   :::|||:::|||   |||   |||::::::||||||
    401 TTATTKGTGGAMC.ATCAAGACAATTCAKGAGAWGAGCAAGAATCCAAAA 449

    273 AlaProArgAspProArgLeuProLysProLysGlyGlnThrValGlySe 289
           |||   |||   ||||||||||||||||||   |||||||||:::||
    450 CCACCGADGGATGCYCGCCTTCCGAAGCCAAAARGGCAGACAGTTGCATC 499

    289 rPheArgLysGlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuG 306
        ||||||||||||||||      ||||||::::::   ::::::   ||||
    500 TTTCAGGAAGGGTCTTGCCAWGCTTCCAAATTCCAATACATCCAGCTTGG 549

    306 lySerLysLeuLysLeuSerTrpLysLeuSerSerIleThrLysSerGlu 322
        ||||||||:::|||||||||   ||||||:::|||||||||||||||:::
    550 GTAGTAAAGTCAAACTATCAWGGAAACTCACGAGCATTACAAAATCAGAT 599

    323 LysGlyGlyTyrHisLeuThrTyrGluThrProGluGlyValValSerLe 339
              ||||||   |||   |||   ||||||:::||||||||||||::
    600 GACAAGGGATATGTTTTGGAGTATTAAACGCCAAAAGGGGTTGTTTCGGT 649

    339 uGlnSerArgSerIleValMetThrValProSerTyrValAlaSerAsnI 356
        :|||::::::|||:::   ||||||:::|||||||||||||||||||||
    650 GCAGGCTAAAAGTTTTAWCATGACTATTCCATCATATGTTGCTAGCAACA 699

    356 leLeuArgProLeuSerValAlaAlaAlaAsp.AlaLeuSerAsnPheTy 372
          |||||||||||||||   :::||||||:::    :::   ::::::||
    700 WTTTGCGTCCACTTTCAAGCKATGCTGCAGAGCCCTATC...AAGATCTA 746

    372 rTyrProProValGlyAlaVal 379
        |   ||||||||:::|||||||
    747 TTAACCACCGGTTGCTGCGGTA 768

seq_name: gb_est99:BG352959

seq_documentation_block:
LOCUS       BG352959       590 bp    mRNA            EST       01-MAR-2001
DEFINITION  sab93f10.y1 Gm-c1040 Glycine max cDNA clone GENOME SYSTEMS CLONE
            ID: Gm-c1040-2588 5' similar to SW:PPOC_ARATH P55826
            PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.
ACCESSION   BG352959
VERSION     BG352959.1  GI:13181619
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
```

```
                    4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                    Tel: 314 286 1800
                    Fax: 314 286 1810
                    Email: est@watson.wustl.edu
                    This clone is available through: Genome Systems, Inc. 4633 World
                    Parkway Circle St. Louis, Missouri 63134 For further information
                    call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
                    427-3324 or contact: clones@genomesystems.com or
                    info@genomesystems.com web site: www.genomesystems.com
                    High quality sequence stop: 481.
FEATURES             Location/Qualifiers
     source          1..590
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1040-2588"
                     /clone_lib="Gm-c1040"
                     /tissue_type="Hypocotyl and Plumule, germinating seeds"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3Pac (Pharmacia); Site_1: EcoRI;
                     Site_2: NotI; This cDNA library was constructed from mRNA
                     isolated from hypocotyl and plumule tissues of seeds
                     germinated for three days of the cultivar Williams 82.
                     Complementary DNA was synthesized from mRNA using a primer
                     consisting  of a poly(dT) sequence with a NotI restriction
                     site. EcoRI adapters were  ligated to the blunt-ended cDNA
                     fragments followed by digestions with EcoRI  and NotI. The
                     cDNA fragments were directionally cloned into the
                     EcoRI-NotI restriction site of the pT7T3-Pac vector. The
                     ligated cDNA fragments were transformed into DH10B host
                     cells (Gibco BRL). This library was constructed by Dr.
                     Randy Shoemaker."
BASE COUNT      156 a    117 c    147 g    170 t
ORIGIN


alignment_scores:
           Quality:  797.00              Length:    191
             Ratio:  4.478                 Gaps:      1
 Percent Similarity:  93.194   Percent Identity: 80.105

alignment_block:
 US-09-508-418-2 x BG352959    ..

 Align seg 1/1  to: BG352959  from: 1  to: 590

     190 GlyProIleGlyLeuArgProSerProProGlyHisGluGluSerValGl 206
         |||   :::|||:::||||||   ||||||||||||||||||||||||||
      15 GGTGCGCTTGGAATTCGGCCTCCTCCTCCAGGTCATGAGGAATCGGTTGA 64

     206 uGlnPheValArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleG 223
         |:::|||||||||||||||||||||   ||||||||||||||||||||||
      65 AGAGTTTGTTCGTCGGAACCTTGGTGATGAGGTTTTTGAACGGTTGATAG 114

     223 luProPheCysSerGlyValTyrValGlyAspProSerLysLeuSerMet 239
         |||||||||||||||||||||||:::||||||||||||||||||||||||
     115 AGCCTTTTTGTTCAGGGGTCTATGCAGGCGATCCTTCAAAATTAAGTATG 164

     240 LysAlaAlaPheGlyLysValTrpLysLeuGluGluThrGlyGlySerIl 256
         |||||||||||||||||||||||||||||||||::::::|||||||||||
     165 AAAGCAGCATTCGGGAAAGTTTGGAAGCTGGAAAAAAATGGTGGTAGCAT 214

     256 eIleGlyGlyThrPheLysAlaIleLysGluArgSerSerThrProLysA 273
         |||||||||||||||||||||||||:::||||||:::::::::   |||
     215 TATTGGTGGAACTTTCAAAGCAATACAAGAGAGAAATGGAGCTTCAAAAC 264

     273 laProArgAspProArgLeuProLysProLysGlyGlnThrValGlySer 289
           ||||||||||||||||||||||||||||||||||||||||||||||||
     265 CACCTCGAGATCCGCGTCTGCCAAAACCAAAAGGTCAGACTGTTGGATCT 314

     290 PheArgLysGlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGl 306
         |||||||||||||||   ||||||||||||||||||||||||||||||||
     315 TTCCGGAAGGGACTTACCATGTTGCCTGATGCAATTTCTGCCAGACTAGG 364
```

```
     306 ySerLysLeuLysLeuSerTrpLysLeuSerSerIleThrLysSerGluL 323
         |:::|||:::|||||||||||||||||||||||||||:::|||   ::::
     365 CAACAAAGTAAAGTTATCTTGGAAGCTTTCAAGTATTAGTAAACTGGATA 414

     323 ysGlyGlyTyrHisLeuThrTyrGluThrProGluGlyValValSerLeu 339
         ::|||   |||   ||||||||||||||||||||||||||||||||||||
     415 GTGGAGAGTACAGTTTGACATATGAAACACCAGAAGGAGTGGTTTCTTTG 464

     340 GlnSerArgSerIleValMetThrValProSerTyrValAlaSerAsnIl 356
         |||   :::::::::|||:::|||:::||||||||||||||||||:::::
     465 CAGTGCAAAACTGTTGTCCTGACCATTCCTTCCTATGTTGCTAGTACATT 514

     356 eLeuArgProLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrT 373
         :|||||||||||||||:::|||||||||||||||      :::|||||||
     515 GCTGCGTCCTCTGTCTGCTGCTGCTGCAGATGCAACTTTCAAGTTTTATT 564

     373 yrPro.ProValGlyAlaVal 379
         ||||| ||||||:::||||||
     565 ACCCTTCCAGTGGCTGCAGTT 585

seq_name: gb_est74:BE455801

seq_documentation_block:
LOCUS       BE455801      580 bp    mRNA            EST       27-FEB-2001
DEFINITION  HVSMEg0015D05f Hordeum vulgare pre-anthesis spike EST library
            HVcDNA0008 (white to yellow anther) Hordeum vulgare cDNA clone
            HVSMEg0015D05f, mRNA sequence.
ACCESSION   BE455801
VERSION     BE455801.2  GI:13154672
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
            ,Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
            ,T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
            Wood,T.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics
  JOURNAL   Unpublished (2000)
COMMENT     On Jul 26, 2000 this sequence version replaced gi:9465836.
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence stop: 499.
FEATURES             Location/Qualifiers
     source          1..580
                     /organism="Hordeum vulgare"
                     /cultivar="Morex"
                     /db_xref="taxon:4513"
                     /clone="HVSMEg0015D05f"
                     /clone_lib="Hordeum vulgare pre-anthesis spike EST library
                     HVcDNA0008 (white to yellow anther)"
                     /tissue_type="pre-anthesis spike"
                     /lab_host="SOLR"
                     /note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1"
BASE COUNT      156 a    124 c    150 g    147 t      3 others
ORIGIN


alignment_scores:
           Quality:  778.00              Length:    192
             Ratio:  4.346                 Gaps:      0
 Percent Similarity:  93.229   Percent Identity: 75.000
```

```
alignment_block:
 US-09-508-418-2 x BE455801   ..

 Align seg 1/1  to: BE455801  from: 1  to: 580

     217 ValPheGluArgLeuIleGluProPheCysSerGlyValTyrValGlyAs 233
         ||||||||||||||||||||||||||||||||||||||||||:::|||||
       3 GTCTTTGAGCGCCTTATCGAACCTTTCTGCTCAGGTGTGTATGCTGGTGA 52

     233 pProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLeuG 250
         |||||||||||||:::|||||||||||||||||||||||||||:::||||
      53 TCCTTCCAAGCTCAATATGAAGGCTGCATTTGGGAAGGTTTGGAGGTTGG 102

     250 luGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysGlu 266
         |||||   ||||||||||||||||||||||||:::|||||||||::::::
     103 AGGAGATTGGGGGTAGTATTATTGGTGGAACCATCAAGGCCATTCAGGAT 152

     267 ArgSerSerThrProLysAlaProArgAspProArgLeuProLysProLy 283
         ::::::::::::||||||   |||||||||||||||||||||   |||||
     153 AAAGGGAAGAACCCCAAACCGCCAAGGGATCCCCGACTTCCGGCACCAAA 202

     283 sGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAspA 300
         |||||||||||||:::||||||:::|||||||||   |||||||||:::|
     203 GGGACAGACGGTGGCATCTTTCAAGAAGGGTCTGGCCATGCTCCCGAATG 252

     300 laIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSer 316
         |||||::::::|||||||||||||||:::||||||||||||||||||:::
     253 CCATCGCATCTAGGTTGGGTAGTAAAGTCAAGCTGTCATGGAAGCTTACG 302

     317 SerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrPr 333
         ||||||||||||:::::::::   ||||||   |||   |||||||||||
     303 AGCATTACAAAGGCGGACAACCAAGGATATGTATTAGGTTATGAAACACC 352

     333 oGluGlyValValSerLeuGlnSerArgSerIleValMetThrValProS 350
         |||||||:::||||||:::|||::::::|||::::::||||||:::||||
     353 AGAAGGACTTGTTTCAGTGCAGGCTAAAAGTGTTATCATGACCATCCCGT 402

     350 erTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAsp 366
         ||||||||||||||:::||||||||||||||||||:::   |||||||||
     403 CATATGTTGCTAGTGATATCTTACGCCCACTTTCAATTGATGCAGCAGAT 452

     367 AlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTy 383
         |||||||||:::||||||||||||||||||:::|||||||||:::|||||
     453 GCACTCTCAAAATTCTATTATCCGCCAGTTGCTGCTGTAACTGTTTCATA 502

     383 rProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLysG 400
         ||||   ||||||||||||   |||      :::|||||||||   :::|
     503 TCCAANAGAAGCTATTAGAAAAGAATGCTTNATTGATGGGGAAGCTCAGG 552

     400 lyPheGlyGlnLeuHisProArgThr 408
         |||||||||||||||||||||||:::
     553 GGTTCGGCCAGCTGCATCCACGTAGC 578

seq_name: gb_est39:AV538014

seq_documentation_block:
LOCUS       AV538014       596 bp    mRNA            EST       06-SEP-2000
DEFINITION  AV538014 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ109a12F 3', mRNA sequence.
ACCESSION   AV538014
VERSION     AV538014.1  GI:8698297
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
```

```
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1..596
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ109a12F"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
                     /tissue_type="roots"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      165 a    135 c    126 g    170 t
ORIGIN


alignment_scores:
           Quality:  739.00             Length:    177
             Ratio:  4.452                Gaps:      0
 Percent Similarity:  93.785   Percent Identity: 79.096

alignment_block:
 US-09-508-418-2 x AV538014/rev   ..

 Align seg 1/1  to reverse of: AV538014  from: 1  to: 596

     372 TyrTyrProProValGlyAlaValThrIleSerTyrProGlnGluAlaIl 388
         |||||||||||||||:::||||||:::||||||||||||:::||||||||
     591 TATTACCCACCAGTTGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAAT 542

     388 eArgAspGluArgLeuValAspGlyGluLeuLysGlyPheGlyGlnLeuH 405
         ||||   |||   |||:::|||||||||||||||||||||||||||||||
     541 CCGAACAGAATGTTTGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGC 492

     405 isProArgThrGlnGlyValGluThrLeuGlyThrIleTyrSerSerSer 421
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     491 ATCCACGCACGCAAGGAGTTGAAACATTAGGAACTATCTACAGCTCCTCA 442

     422 LeuPheProAsnArgAlaProLysGlyArgValLeuLeuLeuAsnTyrIl 438
         |||||||||||||||||||||   ||||||:::|||||||||||||||||
     441 CTCTTTCCAAATCGCGCACCGCCCGGAAGAATTTTGCTGTTGAACTACAT 392

     438 eGlyGlyAlaLysAsnProGluIleLeuSerLysThrGluSerGlnLeuV 455
         |||||||:::   |||      ||||||||||||:::|||::::::||||
     391 TGGCGGGTCTACAAACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAG 342

     455 alGluValValAspArgAspLeuArgLysMetLeuIleLysProLysAla 471
         |||||:::||||||||||||||||||||||||||||||||||||::::::
     341 TGGAAGCAGTTGACAGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCG 292

     472 GlnAspProLeuValValGlyValArgValTrpProGlnAlaIleProGl 488
            |||||||||   :::||||||||||||||||||||||||||||||||
     291 ACCGATCCACTTAAATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCA 242

     488 nPheLeuValGlyHisLeuAspThrLeuSerThrAlaLysAlaAlaMetA 505
         ||||||||||||||||:::|||   |||:::|||||||||::::::::::
     241 GTTTCTAGTTGGTCACTTTGATATCCTTGACACGGCTAAATCATCTCTAA 192

     505 snAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAsnTyrValSerGly 521
         ::::::::|||   ||||||||||||||||||||||||||||||:::|||
     191 CGTCTTCGGGCTACGAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGT 142

     522 ValAlaLeuGlyArgCysValGluGlyAlaTyrGluValAlaSerGluVa 538
         ||||||||||||||||||||||||||||||||||||:::|||   |||||
     141 GTAGCCTTAGGCCGGTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGT 92

     538 lThrGlyPheLeuSerArgTyrAlaTyrLys 548
         |::::::|||:::||||||||||||||||||
      91 CAACAACTTCATGTCACGGTACGCTTACAAG 61
```

```
seq_name: gb_est90:BF624024

seq_documentation_block:
LOCUS       BF624024      510 bp .  mRNA            EST       21-FEB-2001
DEFINITION  HVSMEa0008L16f Hordeum vulgare seedling shoot EST library
            HVcDNA0001 (Cold stress) Hordeum vulgare cDNA clone HVSMEa0008L16f,
            mRNA sequence.
ACCESSION   BF624024
VERSION     BF624024.2  GI:13082749
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 510)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
            ,Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
            ,T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
            Wood,T.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics
  JOURNAL   Unpublished (2000)
COMMENT     On Dec 18, 2000 this sequence version replaced gi:11887758.
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence stop: 505.
FEATURES             Location/Qualifiers
     source          1. .510
                     /organism="Hordeum vulgare"
                     /cultivar="Morex"
                     /db_xref="taxon:4513"
                     /clone="HVSMEa0008L16f"
                     /clone_lib="Hordeum vulgare seedling shoot EST library
                     HVcDNA0001 (Cold stress)"
                     /tissue_type="Seedling shoot"
                     /lab_host="TJC121"
                     /note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1"
BASE COUNT      148 a    110 c    116 g    135 t      1 others
ORIGIN


alignment_scores:
           Quality:  685.00             Length:    168
             Ratio:  4.363                Gaps:      0
 Percent Similarity:  93.452   Percent Identity: 77.381

alignment_block:
 US-09-508-418-2 x BF624024   ..

 Align seg 1/1  to: BF624024  from: 1  to: 510

     320 LysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrProGluGlyVa 336
         |||:::::::::   ||||||   |||   ||||||||||||||||||::
       6 AAGGCGGACAACCAAGGATATGTATTAGGTTATGAAACACCAGAAGGACT 55

     336 lValSerLeuGlnSerArgSerIleValMetThrValProSerTyrValA 353
         :||||||:::|||::::::|||::::::||||||:::|||||||||||||
      56 TGTTTCAGTGCAGGCTAAAAGTGTTATCATGACCATCCCGTCATATGTTG 105

     353 laSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAlaLeuSer 369
         |||||:::||||||||||||||||||:::   ||||||||||||||||||
     106 CTAGTGATATCTTACGCCCACTTTCAATTGATGCAGCAGATGCACTCTCA 155

     370 AsnPheTyrTyrProProValGlyAlaValThrIleSerTyrProGlnGl 386
         :::||||||||||||||||||:::|||||||||:::|||||||||:::||
```

```
     156 AAATTCTATTATCCGCCAGTTGCTGCTGTAACTGTTTCATATCCAAAAGA 205

     386 uAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGlyPheGlyG 403
         ||||||||||   |||   |||:::||||||||||||:::||||||||||
     206 AGCTATTAGAAAAGAATGCTTAATTGATGGGGAGCTCCAGGGTTTCGGCC 255

     403 lnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrIleTyrSer 419
         ||||||||||||||:::|||||||||||||||||||||||||||||||||
     256 AGCTGCATCCACGTAGCCAAGGAGTCGAGACTTTAGGGACAATATATAGC 305

     420 SerSerLeuPheProAsnArgAlaProLysGlyArgValLeuLeuLeuAs 436
         |||||||||||||||||||||||||||   ||||||||||||||||||||
     306 TCTTCTCTCTTTCCCAATCGTGCTCCTGCTGGAAGAGTGTTACTTCTGAA 355

     436 nTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThrGluSerG 453
         |||||||||||||:::   |||      |||:::|||||||||||||||:
     356 CTATATCGGGGGTTCTACAAATACAGGGATCGTCTCCAAGACCGAGAGTG 405

     453 lnLeuValGluValValAspArgAspLeuArgLysMetLeuIleLysPro 469
         ::|||||||||:::||||||||||||||||||||||||||||||:::|||
     406 ACTTAGTAGAAGCTGTTGATCGTGATCTCAGAAAAATGTTGATAAACCCT 455

     470 LysAlaGlnAspProLeuValValGlyValArgValTrpProGlnAlaIl 486
         :::|||   ||||||||||::::::|||||||||||||||||||||||||
     456 AGAGCAGCAGACCCTTTAGCATTAGGGGTCAGAGTGTGGCCACAAGCAAT 505

     486 ePro 487
         ||||
     506 ACCA 509

seq_name: gb_est42:AW119581

seq_documentation_block:
LOCUS       AW119581      498 bp    mRNA            EST       17-JUL-2000
DEFINITION  sd48d01.y1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1016-3002 5' similar to SW:PPOC_ARATH P55826 PROTOPORPHYRINOGEN
            OXIDASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.
ACCESSION   AW119581
VERSION     AW119581.1  GI:6094914
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Insert Length: 1114   Std Error: 0.00
            Seq primer: -40RP from Gibco
            High quality sequence stop: 425.
FEATURES             Location/Qualifiers
     source          1. .498
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
```

```
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1016-3002"
                     /clone_lib="Gm-c1016"
                     /tissue_type="immature flowers of field grown plants"
                     /lab_host="XL10-Gold"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from immature flowers of field grown plants. The cDNA
                     library was prepared using the Stratagene pBluescript II
                     XR library construction kit. Complementary DNA was
                     synthesized from mRNA using a primer consisting of a poly
                     (dT) sequence with a XhoI restriction site. EcoRI adapters
                     were ligated to the blunt-ended cDNA fragments followed by
                     XhoI digestion. The cDNA fragments were directionally
                     cloned into the EcoRI-XhoI restriction site of the
                     pBluescript vector. The ligated cDNA fragments were
                     transformed into XL10-Gold host cells. This library was
                     constructed by Dr. Randy Shoemaker and Dr. John
                     Erpelding."
BASE COUNT      137 a    106 c    111 g    143 t      1 others
ORIGIN


alignment_scores:
           Quality:  674.00            Length:    164
             Ratio:  4.266               Gaps:      0
 Percent Similarity:  96.341   Percent Identity: 82.317

alignment_block:
 US-09-508-418-2 x AW119581   ..

 Align seg 1/1  to: AW119581  from: 1  to: 498

     328 LeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgSerIl 344
         |||||||||||||||||||||||||||||||||||||||   ::::::::
       8 TTGACATATGAAACACCAGAAGGAGTGGTTTCTTTGCAGTGCAAAACTGT 57

     344 eValMetThrValProSerTyrValAlaSerAsnIleLeuArgProLeuS 361
         :|||:::|||:::||||||||||||||||||::::::|||||||||||||
      58 TGTCCTGACCATTCCTTCCTATGTTGCTAGTACATTGCTGCGTCCTCTGT 107

     361 erValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProValGly 377
         ||:::|||||||||||||||||||||:::||||||||||||||||||:::
     108 CTGCTGCTGCTGCAGATGCACTTTCAAAGTTTTATTACCCTCCAGTTGCT 157

     378 AlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgLeuVa 394
         ||||||:::||||||||||||:::||||||||||||:::|||   |||::
     158 GCAGTTTCCATATCCTATCCAAAAGAAGCTATTAGATCAGAATGCTTGAT 207

     394 lAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyV 411
         :|||||||||||||||||||||||||||||||||||||||:::|||||||
     208 AGATGGTGAGTTGAAGGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAG 257

     411 alGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnArgAla 427
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     258 TGGAAACATTAGGAACTATATACAGCTCATCACTATTCCCCAACCGAGCA 307

     428 ProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnPr 444
         |||   ||||||||||||||||||||||||||||||||||||   |||
     308 CCACCTGGAAGGGTTCTACTCTTGAATTACATTGGAGGAGCAACTAATAC 357

     444 oGluIleLeuSerLysThrGluSerGlnLeuValGluValValAspArgA 461
             |||||||||||||||:::|||:::   ||||||:::||||||||||
     358 TGGAATTTTATCGAAGACGGACAGTGAACT.GTGGAAACAGTTGATCGAG 406

     461 spLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuValVal 477
         |||||||||||:::||||||:::|||:::||||||||||||:::||||||
     407 ATTTGAGGAAAATCCTTATAAACCCAAATGCCCAGGATCCATTTGTAGTN 456

     478 GlyValArgValTrpProGlnAlaIleProGlnPheLeuVal 491
         |||||||||:::||||||||||||||||||||||||||||||
     457 GGGGTGAGACTGTGGCCTCAAGCTATTCCACAGTTCTTAGTT 498

seq_name: gb_est69:BE021686
```

```
seq_documentation_block:
LOCUS       BE021686      480 bp    mRNA            EST       21-NOV-2000
DEFINITION  sm61b05.y1 Gm-c1028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1028-8362 5' similar to SW:PPOC_ARATH P55826 PROTOPORPHYRINOGEN
            OXIDASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.
ACCESSION   BE021686
VERSION     BE021686.1  GI:8284127
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Insert Length: 1299   Std Error: 0.00
            High quality sequence stop: 394.
FEATURES             Location/Qualifiers
     source          1..480
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1028-8362"
                     /clone_lib="Gm-c1028"
                     /tissue_type="roots of 'Supernod' plants"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; The mRNA was isolated from roots of Glycine max
                     'Supernod' plants generously donated by Dr. Gary Stacey.
                     The seedlings were innoculated with Bradyrhizobium
                     japonicus, strain USDA110 priot to harvest. Stratagene's
                     cDNA synthesis Kit (catalog number 200401) was used to
                     synthesize the cDNA.  First-strand synthesis was performed
                     with 5-methyl dCTP, hence the ligated cDNA was
                     hemimethylated.  A modification of Stratagene's
                     first-strand synthesis primer was used.  An 'anchor'
                     nucleotide (V=A,C, or G) was added to the 3' end of the
                     primer [GAGAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18V] to anchor
                     the primer at the 5' end of the poly(A) tract.  After
                     second-strand synthesis, the cDNA ends were filled in with
                     cloned Pfu DNA polymerase, ligated to EcoRI adapters and
                     subsequently phosphorylated.  The XhoI site within the
                     first-strand synthesis primer was then restricted by
                     digestion with XhoI; all XhoI sites in the cDNA would be
                     protected by their hemimethylated status.  The cDNA
                     constructs were size-fractionated with a 500bp cutoff,
                     using GibcoBRL Life Technologies' cDNA Size Fractionation
                     column.  The column eluent was then ligated into
                     Stratagene's pBluescript II XR Predigested vector
                     (pBluescript II SK(+) that has been digested with EcoRI
                     and XhoI, and phosporylated by Stratagene).  Both the
                     white and blue colonies appear to contain recombinant
                     plasmids with cDNA inserts, based on size (n=25).  This
                     library was constructed by Dr. Paul Keim and  Dr. Virginia
```

```
                Coryell."
BASE COUNT      138 a     98 c    104 g    137 t      3 others
ORIGIN


alignment_scores:
            Quality:  629.00              Length:    158
              Ratio:  4.338                 Gaps:      0
 Percent Similarity:  91.772    Percent Identity: 77.848

alignment_block:
 US-09-508-418-2 x BE021686   ..

 Align seg 1/1  to: BE021686  from: 1  to: 480

     297 LeuProAspAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTr 313
         ||||||||||||||||||||||||||||||:::|||:::|||||||||||
       7 TTGCCTGATGCAATTTCTGCCAGACTAGGCAACAAAGTAAAGTTATCTTG 56

     313 pLysLeuSerSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrT 330
         ||||||||||||||||:::|||   ::::::|||   |||   |||||||
      57 GAAGCTTTCAAGTATTAGTAAACTGGATAGTGGAGAGTACAGTTTGACAT 106

     330 yrGluThrProGluGlyValValSerLeuGlnSerArgSerIleValMet 346
         ||||||||||||||||||||||||||||||||   :::::::::|||:::
     107 ATGAAACACCAGAAGGAGTGGTTTCTTTGCAGTGCAAAACTGTTGTCCTG 156

     347 ThrValProSerTyrValAlaSerAsnIleLeuArgProLeuSerValAl 363
         |||:::||||||||||||||||||::::::|||||||||||||||:::||
     157 ACCATTCCTTCCTATGTTGCTAGTACATTGCTGCGTCCTCTGTCTGCTGC 206

     363 aAlaAlaAspAlaLeuSerAsnPheTyrTyrProProValGlyAlaValT 380
         |||||||||||||||||||:::||||||||||||||||||:::||||||:
     207 TGCTGCAGATGCACTTTCAAAGTTTATTACCCTCCAGTTGCTGCAGTTT 256

     380 hrIleSerTyrProGlnGluAlaIleArgAspGluArgLeuValAspGly 396
         ::||||||||||||:::||||||||||||:::|||   |||:::||||||
     257 CCATATCCTATCCAAAAGAAGCTATTAGATCAGAATGCTTGATAGATGGT 306

     397 GluLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluTh 413
         |||||||||||||||||||||||||||||||||:::||||||||||||||
     307 GAGTTGAAGGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAAC 356

     413 rLeuGlyThrIleTyrSerSerSerLeuPheProAsnArgAlaProLysG 430
         ||||   |||||||||||||||||||||||||||||||||||||||
     357 ATTANGAACTATATACAGCTCATCACTATTCCCCAACCGAGCACCACCTG 406

     430 lyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIle 446
           ||||||||||||   ||||||||||||||||||   |||      |||
     407 AAAGGGTTCTACTCTNGAATTACATTGGAGGAGCAACTAATACTGGAATT 456

     447 LeuSerLysThrGluSerGlnLeu 454
            |||||||||:::|||:::|||
     457 NTATCAAAGACGGACAGTGAACTT 480

 seq_name: gb_est99:BG355389

 seq_documentation_block:
 LOCUS       BG355389       479 bp    mRNA            EST       06-MAR-2001
 DEFINITION  947043H05.y2 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA
             sequence.
 ACCESSION   BG355389
 VERSION     BG355389.1  GI:13237375
 KEYWORDS    EST.
 SOURCE      Zea mays.
   ORGANISM  Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
             clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE   1  (bases 1 to 479)
   AUTHORS   Walbot,V.
  ,TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
             University
```

```
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 947043  row: H  column: 05.
FEATURES             Location/Qualifiers
     source          1..479
                     /organism="Zea mays"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone_lib="947 - 2 week shoot from Barkan lab"
                     /tissue_type="leaf and stem, including leaf base"
                     /dev_stage="2 week old seedling (3 leaves)"
                     /lab_host="XL1-Blue"
                     /note="Organ: shoot; Vector: Lambda ZAP (pBlueScript SK-);
                     Site_1: EcoRI; Site_2: XhoI; Directionally cloned using
                     Stratagene's UniZap XR cDNA cloning kit with  the 5' end
                     at the EcoRI site. The library represents 8 x 10e5
                     independent recombinant phage. The plants were greenhouse
                     grown."
BASE COUNT      122 a    108 c    126 g    123 t
ORIGIN


alignment_scores:
            Quality:  621.00              Length:    151
              Ratio:  4.500                 Gaps:      0
 Percent Similarity:  91.391    Percent Identity: 79.470

alignment_block:
 US-09-508-418-2 x BG355389   ..

 Align seg 1/1  to: BG355389  from: 1  to: 479

     398 LeuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLe 414
         |||:::||||||||||||||||||||||||:::|||||||||||||||||
       2 CTCCAGGGCTTTGGCCAGTTGCATCCACGTAGTCAAGGAGTTGAGACATT 51

     414 uGlyThrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyA 431
         |||||||||||||||||||||||||||||||||||||||||||   ||||
      52 AGGAACAATATACAGTTCCTCACTCTTTCCAAATCGTGCTCCTGACGGTA 101

     431 rgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeu 447
         ||||||||||||||||||||||||||||||||   |||      |||:::
     102 GGGTGTTACTTCTAAACTACATAGGAGGTGCTACAAACACAGGAATTGTT 151

     448 SerLysThrGluSerGlnLeuValGluValValAspArgAspLeuArgLy 464
         |||||||||||||||:::|||||||||   ||||||||||||||||||||
     152 TCCAAGACTGAAAGTGAGCTGGTCGAATCAGTTGACCGTGACCTCCGAAA 201

     464 sMetLeuIleLysProLysAlaGlnAspProLeuValValGlyValArgV 481
         ||||||||||:::      |||   ||||||||||||:::||||||||||
     202 AATGCTTATAAATTCTACAGCAGTGGACCCTTTAGTCCTTGGTGTTCGAG 251

     481 alTrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeu 497
         ||||||||||||||||||||||||||||||||||||||||||||   |||
     252 TTTGGCCACAAGCCATACCTCAGTTCCTGGTAGGACATCTTGATCTTCTG 301

     498 SerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLe 514
         ::::::||||||||||||::::::   :::|||   :::|||||||||||
     302 GAAGCCGCAAAAGCTGCCCTGGACCGAGGTGGCTACGATGGGCTGTTCCT 351

     514 uGlyGlyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyA 531
         ||||||||||||||||:::|||||||||||||||||||||||||||||||
     352 AGGAGGGAACTATGTTGCAGGAGTTGCCCTGGGCAGATGCGTTGAGGGCG 401

     531 laTyrGluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyr 547
         ||||||||   ||||||:::::::::   ||||||::::::|||||||||
     402 CGTATGAAAGTGCCTCGCAAATATCTGACTTCTTGACCAAGTATGCCTAC 451
```

```
    548 Lys 548
        |||
    452 AAG 454

seq_name: gb_est39:AV540305

seq_documentation_block:
LOCUS       AV540305      519 bp    mRNA            EST       06-SEP-2000
DEFINITION  AV540305 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ148e10F 3', mRNA sequence.
ACCESSION   AV540305
VERSION     AV540305.1  GI:8702063
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 519)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .519
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ148e10F"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
                     /tissue_type="roots"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      150 a    116 c    109 g    144 t
ORIGIN


alignment_scores:
          Quality:  605.00             Length:    146
            Ratio:  4.416                Gaps:      0
 Percent Similarity:  93.836   Percent Identity: 78.767

alignment_block:
 US-09-508-418-2 x AV540305/rev    ..

 Align seg 1/1  to reverse of: AV540305  from: 1  to: 519

    403 GlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrIleTyrSe 419
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    519 CAATTGCATCCGCGGACGCAAGGAGTTGAAACATTAGGAACTATCTACAG 470

    419 rSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuLeuLeuA 436
        ||||||||||||||||||||||||||||   ||||||:::||||||||||
    469 CTCCTCACTCTTTCCAAATCGCGCACCGCCCGGAAGAATTTTGCTGTTGA 420

    436 snTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThrGluSer 452
        ||||||||||||||:::   |||      ||||||||||||:::|||:::
    419 ACTACATTGGCGGGTCTACAAACACCGGAATTCTGTCCAAGTCTGAAGGT 370

    453 GlnLeuValGluValValAspArgAspLeuArgLysMetLeuIleLysPr 469
        :::|||||||||:::|||||||||||||||||||||||||||||||||||
    369 GAGTTAGTGGAAGCAGTTGACAGAGATTTGAGGAAAATGCTAATTAAGCC 320

    469 oLysAlaGlnAspProLeuValValGlyValArgValTrpProGlnAlaI 486
        |::::::   |||||||||   :::|||||||||||||||||||||||||
```

```
    319 TAATTCGACCGATCCACTTAAATTAGGAGTTAGGGTATGGCCTCAAGCCA 270

    486 leProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAlaLysAla 502
        |||||||||||||||||||||||:::|||   |||:::|||||||||:::
    269 TTCCTCAGTTTCTAGTTGGTCACTTTGATATCCTTGACACGGCTAAATCA 220

    503 AlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAsnTyrVa 519
        :::::::::::::::|||   |||||||||||||||||||||||||||||
    219 TCTCTAACGTCTTCGGGCTACGAAGGGCTATTTTGGGTGGCAATTACGT 170

    519 lSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluValAlaS 536
        |:::|||||||||||||||||||||||||||||||||||||||:::|||
    169 CGCTGGTGTAGCCTTAGGCCGGTGTGTAGAAGGCGCATATGAAACCGCGA 120

    536 erGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
          ||||||::::::|||:::||||||||||||||||||
    119 TTGAGGTCAACAACTTCATGTCACGGTACGCTTACAAG 82

seq_name: gb_est24:AI748531

seq_documentation_block:
LOCUS       AI748531      457 bp    mRNA            EST       17-JUL-2000
DEFINITION  sb54d01.y1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1016-314 5' similar to SW:PPOC_ARATH P55826 PROTOPORPHYRINOGEN
            OXIDASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.
ACCESSION   AI748531
VERSION     AI748531.1  GI:5126795
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Insert Length: 1096   Std Error: 0.00
            High quality sequence stop: 401.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1016-314"
                     /clone_lib="Gm-c1016"
                     /tissue_type="immature flowers of field grown plants"
                     /lab_host="XL10-Gold"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from immature flowers of field grown plants. The cDNA
                     library was prepared using the Stratagene pBluescript II
                     XR library construction kit. Complementary DNA was
                     synthesized from mRNA using a primer consisting of a poly
                     (dT) sequence with a XhoI restriction site. EcoRI adapters
                     were ligated to the blunt-ended cDNA fragments followed by
                     XhoI digestion. The cDNA fragments were directionally
```

```
                cloned into the EcoRI-XhoI restriction site of the
                pBluescript vector. The ligated cDNA fragments were
                transformed into XL10-Gold host cells. This library was
                constructed by Dr. Randy Shoemaker and Dr. John
                Erpelding."
BASE COUNT      124 a     96 c    106 g    131 t
ORIGIN


alignment_scores:
           Quality:  593.00              Length:    152
             Ratio:  4.090                 Gaps:      1
 Percent Similarity:  95.395   Percent Identity: 80.263

alignment_block:
 US-09-508-418-2 x AI748531   ..

 Align seg 1/1  to: AI748531  from: 1  to: 457

     335 GlyValValSerLeuGlnSerArgSerIleValMetThrValProSerTy 351
         ||||||||||||||||||   :::::::::|||:::|||:::||||||||
       3 GGAGTGGTTTCTTTGCAGTGCAAAACTGTTGTCCTGACCATTCCTTCCTA 52

     351 rValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAlaL 368
         ||||||||||::::::|||||||||||||||:::||||||||||||||||
      53 TGTTGCTAGTACATTGCTGCGTCCTCTGTCTGCTGCTGCAGATGCAC 102

     368 euSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrPro 384
         |||||:::||||||||||||||||||:::||||||:::||||||||||||
     103 TTTCAAAGTTTTATTACCCTCCAGTTGCTGCAGTTTCCATATCCTATCCA 152

     385 GlnGluAlaIleArgAspGluArgLeuValAsp.GlyGluLeuLysGlyP 401
         :::||||||||||||:::|||   |||:::||| ||||||||||||||||
     153 AAAGAAGCTATTAGATCAGAATGCTTGATAGATGGGTGAGTTGAAGGGGT 202

     401 heGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrIle 417
         ||||||||||||||||||||:::|||||||||||||||||||||||||||
     203 TTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAACATTAGGAACTATA 252

     418 TyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuLe 434
         |||||||||||||||||||||||||||||||||   ||||||||||||||
     253 TACAGCTCATCACTATTCCCCAACCGAGCACCACCTGGAAGGGTTCTACT 302

     434 uLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThrG 451
         ||||||||||||||||||||||   |||      |||||||||||||||:
     303 CTTGAATTACATTGGAGGAGCAACTAATACTGGAATTTTATCGAAGACGG 352

     451 luSerGlnLeuValGluValValAspArgAspLeuArgLysMetLeuIle 467
         ::|||:::   ||||||:::|||||||||||||||||||||:::||||||
     353 ACAGTGAACT.GTGGAAACAGTTGATCGAGATTTGAGGAAAATCCTTATA 401

     468 LysProLysAlaGlnAspProLeuValValGlyValArgValTrpProGl 484
         :::|||:::||||||||||||:::|||||||||||||||:::||||||||
     402 AACCCAAATGCCCAGGATCCATTTGTAGTGGGGGTGAGACTGTGGCCTCA 451

     484 nAla 485
         ||||
     452 AGCT 455

 seq_name: gb_est97:BG133906

 seq_documentation_block:
 LOCUS       BG133906      440 bp    mRNA            EST       31-JAN-2001
 DEFINITION  EST466798 tomato crown gall Lycopersicon esculentum cDNA clone
             cTOE14B22 5' sequence, mRNA sequence.
 ACCESSION   BG133906
 VERSION     BG133906.1  GI:12634094
 KEYWORDS    EST.
 SOURCE      tomato.
   ORGANISM  Lycopersicon esculentum
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
```

```
            Lycopersicon.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   van der Hoeven,R., Sun,H., Cho,J., Utterback,T., Hansen,C., Ronning
            ,C. and Tanksley,S.
  TITLE     Generation of ESTs from tomato crown gall tissue
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES             Location/Qualifiers
     source          1. .440
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOE14B22"
                     /clone_lib="tomato crown gall"
                     /tissue_type="crown gall"
                     /dev_stage="crown galls from full-grown plants (8 wks old
                     )"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; Four wk old greenhouse plants were stab inoculated
                     on stem with Agrobacterium tumefaciens C58 (Dr. T.J. Burr,
                     Cornell U.). Galls were allowed to develop for another 4
                     wks, when gall tissue was frozen in  liquid nitrogen."
BASE COUNT     104 a    105 c    109 g    122 t
ORIGIN


alignment_scores:
           Quality:  589.00              Length:    145
             Ratio:  4.675                 Gaps:      2
 Percent Similarity:  86.897   Percent Identity: 77.931

alignment_block:
 US-09-508-418-2 x BG133906   ..

 Align seg 1/1  to: BG133906  from: 1  to: 440

       1 MetThrThrThrProIleAlaAsnHisProAsnIlePheThrHisGln.. 16
         ||||||||||||   ::::::|||||||||:::||||||||||||:::
       5 ATGACAACAACGGCCGTCGTCAACCATCCTAGCATTTTCACTCACCGGTC 54

      17 .........................SerSerSerSerPro...LeuAlaP 24
                                  |||||||||||||||   :::   |
      55 GCCGCTGCCGTCGCCGTCCTCCTCCTCATCCTCATCGCCGTCATTTTTAT 104

      24 heLeuAsnArgThrSerPheIleProPheSerSerIleSerLysArgAsn 40
         ||||||||||||||:::|||||||||:::   |||   |||||||||:::
     105 TTTTAAATCGTACGAATTTTATTCCATACTTTTCCACCTCCAAGCGCAGT 154

      41 SerValAsnCysAsnGlyTrpArgThrArgCysSerValAlaLysAspTy 57
         |||||||||||||||||||||||||||||||||||||||||||||:::||
     155 AGTGTCAATTGCAATGGCTGGAGAACACGGTGTTCCGTTGCGAAGAATTA 204

      57 rThrValProSerSerAlaValAspGlyGlyProAlaAlaGluLeuAspC 74
         ||||||||||   |||   |||||||||:::         ||||||||||
     205 TACAGTTCCTCCCTCAGAAGTTGACGGTAATCAGTTACCGGAGCTGGATT 254

      74 ysValIleValGlyAlaGlyIleSerGlyLeuCysIleAlaGlnValMet 90
         |||||:::|||||||||||||||||||||||||||||||||:::|||:::
     255 GTGTGGTAGTCGGAGCAGGAATTAGTGGTCTCTGCATTGCTAAGGTGATA 304

      91 SerAlaAsnTyrProAsnLeuMetValThrGluAlaArgAspArgAlaGl 107
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     305 TCGGCTAATTATCCCAATTTGATGGTGACGGAGGCGAGGGATCGTGCCGG 354

     107 yGlyAsnIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyP 124
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     355 TGGAAACATAACGACGGTGGAAAGAGATGGATACTTATGGGAAGAAGGTC 404
```

```
    124 roAsnSerPheGlnProSerAspProMetLeuThr 135
        |||||||||||||||||||||||||||||||||||
    405 CTAACAGTTTCCAGCCTTCGGATCCTATGTTGACT 439

seq_name: gb_est73:BE326089

seq_documentation_block:
LOCUS       BE326089      590 bp    mRNA            EST       21-DEC-2000
DEFINITION  NF084G10ST1F1081 Developing stem Medicago truncatula cDNA clone
            NF084G10ST 5', mRNA sequence.
ACCESSION   BE326089
VERSION     BE326089.2  GI:11936674
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
            ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
            ,R.A.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula stem library
  JOURNAL   Unpublished (2000)
COMMENT     On Jul 14, 2000 this sequence version replaced gi:9199866.
            Contact: Dixon RA
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7302
            Fax: 580 221 7380
            Email: radixon@noble.org
            Medicago Genome Initiative accession: MGI:S:26673
            Insert Length: 703   Std Error: 0.00
            Plate: 084  row: G  column: 10
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF084G10ST"
                     /clone_lib="Developing stem"
                     /tissue_type="stem"
                     /dev_stage="Pooled developmental"
                     /note="Vector: Lambda Zap; Contains a mixture of
                     internodal stem segments"
BASE COUNT     130 a    153 c    131 g    176 t
ORIGIN


alignment_scores:
           Quality:  552.50             Length:    198
             Ratio:  3.635                Gaps:      7
 Percent Similarity:  76.768    Percent Identity: 59.596

alignment_block:
 US-09-508-418-2 x BE326089   ..

 Align seg 1/1  to: BE326089  from: 1  to: 590

      8 AsnHisProAsnIlePheThrHisGlnSerSerSerSerProLeuAlaPh 24
        :::||||||   :::   ||||||::::::   |||   ||||||
     21 TCTCACCCT...CTTTCCACCCACTCAAACCCTTCTCCGCCCTTA..... 62

     24 eLeuAsnArgThrSer.PheIleProPheSerSer............... 35
                     |||       ||||||   |||
     63 .............TCTCCATTCCCCTTTCCCCTCCCCCAATTCAAAATCA 99

     36 ......IleSerLysArgAsnSerValAsnCysAsnGlyTrpArgThrAr 50
              |||   :::||||||      |||            |||   ||
    100 AAATTCATTCGCAATCGCAATCACCATAAC...CCCATTCTCCGTCTCCG 146

     50 gCysSerValAlaLysAspTyrThrVal............ProSerSerA 63
        |||||||:::|||::::::   |||:::               :::|||
    147 TTGCTCCATTGCGGAGGAATCCACCGCATCTCCGACTAAAACCACCTCAA 196

     63 laValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAla 79
             ::::::|||   ::::::::::::|||||||||:::||||||:::
    197 AATCAAATTCCGGAGAGTCCTCATCCGTTGATTGTGTAGTTGTTGGTGGT 246

     80 GlyIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
        |||||||||||||||||||||||||||::::::|||:::::::::
    247 GGCATCAGTGGCCTTTGCATCGCTCAGGCTCTTTCTACAAAGCATGCTCA 296

     95 .......ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyA 109
               ||||||::::::||||||||||||:::||||||:::|||||||
    297 TGGTGTCCCCAATGTTGTTGTTACAGAGGCGAAAGACCGAGTTGGTGGGA 346

    109 snIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsn 125
        |||||   ||||||||||||||||||:::|||||||||||||||||||||
    347 ATATTATTACTGTGGAGAGAGATGGGTTTCTTTGGGAAGAAGGTCCTAAT 396

    126 SerPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLe 142
        |||||||||||||||||||||||||||||||||:::||||||   |||||
    397 AGTTTTCAACCTTCTGATCCTATGCTCACCATGGTGGTGGACAGTGGTTT 446

    142 uLysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuT 159
        ||||||||||||||||||||||||||||:::|||||||||||||||||||
    447 GAAGGACGATCTTGTTTTGGGTGATCCTGATGCACCTAGATTTGTGTTGT 496

    159 rpLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhe 175
        ||:::|||||||||||||||||||||:::|||   |||||||||||||||
    497 GGAATGGGAAATTGAGGCCGGTGCCGGGAAAACCAACTGATTTGCCTTTC 546

    176 PheAspLeuMetSerIleProGlyLysLeuArgAlaGlyPhe 189
        ||||||||||||||||||   ||||||::::::|||||||||
    547 TTTGACTTAATGAGTATTGGTGGCAAGATCAAGGCTGGCTTT 588

seq_name: gb_est51:AW755783

seq_documentation_block:
LOCUS       AW755783      451 bp    mRNA            EST       21-NOV-2000
DEFINITION  s109c11.y1 Gm-c1036 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1036-813 5' similar to SW:PPOC_ARATH P55826 PROTOPORPHYRINOGEN
            OXIDASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.
ACCESSION   AW755783
VERSION     AW755783.1  GI:7685135
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Insert Length: 894   Std Error: 0.00
```

```
            High quality sequence stop: 298.
FEATURES             Location/Qualifiers
     source          1. .451
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1036-813"
                     /clone_lib="Gm-c1036"
                     /tissue_type="somatic embryos cultured on MSD 20"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; This
                     cDNA library was constructed from mRNA isolated from
                     somatic  embryos (age ranging from 2 months to 9 months)
                     cultured on MSD 20. The library was prepared using the
                     Life Technologies pSuperScript cDNA library construction
                     kit. Complementary DNA was synthesized from mRNA using a
                     poly (dT) sequence with a NotI restrictions site. SalI
                     linkers adapters were ligated to the blunt-ended cDNA
                     fragments followed by NotI digestion. The cDNA fragments
                     were directionally cloned into the NotI-SalI restriction
                     site of the pSPORT1 vector. The ligated cDNA fragments
                     were transformed into E.coli ElectroMax DH10B host cells.
                     This library was constructed in the laboratory of Dr. Lila
                     Vodkin by Anu Khanna at the University of Illinois at
                     Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"
BASE COUNT      125 a     88 c    111 g    127 t
ORIGIN


alignment_scores:
          Quality:  543.00            Length:    146
            Ratio:  4.022               Gaps:      0
 Percent Similarity:  92.466   Percent Identity: 75.342

alignment_block:
 US-09-508-418-2 x AW755783   ..

 Align seg 1/1  to: AW755783  from: 1  to: 451

     382 SerTyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLe 398
         |||   |||:::||||||||||||:::|||   |||:::|||||||||||
      11 TCCTGTCCAAAAGAAGCTATTAGATCAGAATGCTTGATAGATGGTGAGTT 60

     398 uLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuG 415
         ||||||||||||||||||||||||||||:::|||||||||||||||||||
      61 GAAGGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAACATTAG 110

     415 lyThrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArg 431
         |||||||||||||||||||||||||||||||||||||||||   ||||||
     111 GAACTATATACAGCTCATCACTATTCCCCAACCGAGCACCACCTGGAAGG 160

     432 ValLeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSe 448
         ||||||||||||||||||||||||||||||   |||      ||||||||
     161 GTTCTACTCTTGAATTACATTGGAGGAGCAACTAATACTGGAATTTTATC 210

     448 rLysThrGluSerGlnLeuValGluValValAspArgAspLeuArgLysM 465
         |||||||:::|||:::|||||||||::::::||||||||||||||||||:
     211 AAAGACGGACAGTGAACTTGTGGAAACAGCTGATCGAGATTTGAGGAAAA 260

     465 etLeuIleLysProLysAlaGlnAspProLeuValValGlyValArgVal 481
         ::||||||:::|||:::||||||||||||:::|||||||||||||||:::
     261 TCCTTATAAACCCAAATGCCCAGGATCCATTTGTAGTGGGGGTGAGACTG 310

     482 TrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSe 498
         |||||||||||||||||||||:::|||   ||||||:::|||   |||::
     311 TGGCCTCAAGCTATTCCACAGCTCTTATTTGGCCATATTGATCTTCTAGA 360

     498 rThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuG 515
         ::::|||||||||:::::::::::::::|||:::|||||||||||||||
     361 TGTGGCTAAAGCTTCTATCAGAAATACTGGGTTTGAAGGGCTCTTTCTTT 409

     515 lyGlyAsnTyrValSerGlyValAlaLeuGlyArgCys 527
           |||:::||||||||||||   :::|||   ||||||
     410 G.GGACTTATGTGTCTGGCGGTTGCTTATGACGATGC 447
```

```
                        GenCore version 4.5
               Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          July  3, 2001, 10:33:02 ; Search time 100.76 Seconds
                                           (without alignments)
                                           719.563 Million cell updates/sec

Title:           US-09-508-418-2
Perfect score:   2844
Sequence:        1 MTTTPIANHPNIFTHQSSSS..........EGAYEVASEVTGFLSRYAYK 548

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:      425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                   %
Result           Query
  No.   Score    Match Length DB  ID                        Description
-----------------------------------------------------------------------------
    1     2832    99.6    548  10  Q9SBI4                  Q9sbi4 nicotiana t
    2     2824    99.3    548  10  Q9SAX9                  Q9sax9 nicotiana t
    3   2565.5    90.2    557  10  O64384                  O64384 solanum tub
    4     2061    72.5    545  10  Q9ZSI5                  Q9zsi5 arabidopsis
    5   2004.5    70.5    555  10  Q9SPL6                  Q9spl6 cichorium i
    6   1989.5    70.0    535  10  Q9M629                  Q9m629 zea mays (m
    7   1966.5    69.1    562  10  Q9LRI8                  Q9lri8 spinacia ol
    8   1530.5    53.8    563  10  Q9ZTA7                  Q9zta7 chlamydomon
    9   1527.5    53.7    563  10  Q9T0P9                  Q9t0p9 chlamydomon
   10    522.5    18.4    502  10  Q9SLW5                  Q9slw5 glycine max
   11      515    18.1    436  2  O67814                  O67814 aquifex aeo
   12      459    16.1    544  10  Q9FYV8                  Q9fyv8 zea mays (m
   13      456    16.0    501  10  Q9LYA7                  Q9lya7 arabidopsis
   14      456    16.0    547  10  Q9FMS9                  Q9fms9 arabidopsis
   15      424    14.9    404  10  O64385                  O64385 solanum tub
   16    407.5    14.3    467  2  Q9KDK8                  Q9kdk8 bacillus ha
   17    363.5    12.8    462  2  Q9RV99                  Q9rv99 deinococcus
   18      317    11.1    468  2  O69829                  O69829 streptomyce
   19    290.5    10.2    391  10  Q9FW53                  Q9fw53 oryza sativ
   20      254     8.9    435  2  Q9JS40                  Q9js40 chlamydia p
   21.     250     8.8    475  5  Q9U8D0                  Q9u8d0 drosophila
   22      249     8.8    435  2  Q9Z718                  Q9z718 chlamydia p
   23      247     8.7    475  5  Q9VC52                  Q9vc52 drosophila
   24    236.5     8.3     63  10  Q9M5P0                  Q9m5p0 nicotiana t
   25    231.5     8.1    424  2  Q9PLH9                  Q9plh9 chlamydia m
   26      211     7.4    424  2  O84750                  O84750 chlamydia t
   27    174.5     6.1    421  2  Q9RSV7                  Q9rsv7 deinococcus
   28    169.5     6.0    441  2  Q9RTN3                  Q9rtn3 deinococcus
   29      162     5.7    524  13  Q9I8A7                  Q9i8a7 scomber jap
   30    153.5     5.4    416  2  Q9RI54                  Q9ri54 streptomyce
   31      148     5.2    440  1  O58162                  O58162 pyrococcus
   32      146     5.1    430  1  Q9HRU3                  Q9hru3 halobacteri
   33      144     5.1    488  10  Q9LYT1                  Q9lyt1 arabidopsis
   34      143     5.0    490  10  Q9SKX5                  Q9skx5 arabidopsis
   35    136.5     4.8    467  2  Q9S2P5                  Q9s2p5 streptomyce
   36      130     4.6    890  5  Q9VW97                  Q9vw97 drosophila
   37      127     4.5    574  10  Q9FVR9                  Q9fvr9 arabidopsis
   38      126     4.4    489  2  Q9KHE1                  Q9khe1 streptomyce
   39      126     4.4    499  2  Q9K566                  Q9k566 mycobacteri
   40      125     4.4    578  10  Q9FVS7                  Q9fvs7 arabidopsis
   41    121.5     4.3    443  2  O82865                  O82865 micrococcus
   42    121.5     4.3    479  2  Q9R6X5                  Q9r6x5 anabaena sp
   43      121     4.3    404  1  O28536                  O28536 archaeoglob
   44    119.5     4.2    472  2  Q9ZBG4                  Q9zbg4 streptomyce
   45    119.5     4.2    570  10  O65813                  O65813 haematococc


                                ALIGNMENTS


RESULT   1
Q9SBI4
ID   Q9SBI4      PRELIMINARY;      PRT;    548 AA.
AC   Q9SBI4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE PX-1 (EC 1.3.3.4).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. SR1;
RA   Horikoshi M., Mametsuka K., Hirooka T.;
RT   "The Molecular Basis of Photobleaching Herbicide Resistance in
RT   Tobacco.";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF044128; AAD02290.1; -.
KW   Oxidoreductase.
SQ   SEQUENCE   548 AA;  59241 MW;  00E1A106F299E4AF CRC64;


  Query Match            99.6%;  Score 2832;  DB 10;  Length 548;
  Best Local Similarity  99.6%;  Pred. No. 1.9e-200;
  Matches  546;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60

Qy      61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120

Qy     121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMS 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMS 180

Qy     181 IPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMK 240
```

```
          |||||||||| ||||||||||||||||||||||||||||||||||||||| |||||||||
Db    181 IPGKLRAGFGAIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYAGDPSKLSMK 240

Qy    241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300

Qy    301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360

Qy    361 SVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 SVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420

Qy    421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480

Qy    481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540

Qy    541 FLSRYAYK 548
          ||||||||
Db    541 FLSRYAYK 548


RESULT   2
Q9SAX9
ID   Q9SAX9      PRELIMINARY;      PRT;   548 AA.
AC   Q9SAX9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PLASTIDAL PROTOPORPHYRINOGEN OXIDASE.
GN   NTPPOX1.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SAMSUN NN;
RX   PubMed=9808719;
RA   Watanabe N., Che F., Iwano M., Takayama S., Nakano T., Yoshida S.,
RA   Isogai A.;
RT   "Molecular characterization of photomixotrophic cultured tobacco cells
RT   resistant to protoporphyrinogen oxidase-inhibiting herbicides.";
RL   Plant Physiol. 118:751-758(1998).
DR   EMBL; AB020501; BAA34713.1; -.
DR   InterPro; IPR000894; -.
DR   ProDom; PD000290; -; 1.
SQ   SEQUENCE   548 AA;  59215 MW;  AB51A513FD3CE7BC CRC64;

  Query Match             99.3%;  Score 2824;  DB 10;  Length 548;
  Best Local Similarity   99.5%;  Pred. No. 7.6e-200;
  Matches  545;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60

Qy     61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120

Qy    121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMS 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLAFFDLMS 180
```

```
Qy    181 IPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMK 240
          |||||||||| ||||||||||||||||||||||||||||||||||||||| |||||||||
Db    181 IPGKLRAGFGAIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYAGDPSKLSMK 240

Qy    241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300

Qy    301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360

Qy    361 SVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 SVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420

Qy    421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480

Qy    481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540

Qy    541 FLSRYAYK 548
          ||||||||
Db    541 FLSRYAYK 548


RESULT   3
O64384
ID   O64384      PRELIMINARY;      PRT;   557 AA.
AC   O64384;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4).
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4113;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. BINTJE;
RA   Johnston D.J., Droz E., Rochaix J.D., Malnoe P.C.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O.
DR   EMBL; AJ225107; CAA12400.1; -.
DR   Mendel; 29307; Soltu;3076;29307.
KW   Oxidoreductase.
SQ   SEQUENCE   557 AA;  60463 MW;  AE2B15B17E0B89F8 CRC64;

  Query Match             90.2%;  Score 2565.5;  DB 10;  Length 557;
  Best Local Similarity   88.9%;  Pred. No. 8.7e-181;
  Matches  495;  Conservative   21;  Mismatches   32;  Indels    9;  Gaps    2;

Qy      1 MTTTPIANHPNIFTHQ--------SSSSP-LAFLNRTSFIPFSSISKRNSVNCNGWRTRC 51
          |||| :||||:||||:        |||||   |||||:|||: | |||||||||||||||
Db      1 MTTTAVANHPSIFTHRSPLPSPSSSSSSPSFLFLNRTNFIPYFSTSKRNSVNCNGWRTRC 60

Qy     52 SVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNIT 111
          ||||||||| | |||    |||||:|||||||||||:|:|||||||||||||||||||||
Db     61 SVAKDYTVPPSEVDGNQFPELDCVVVGAGISGLCIAKVISANYPNLMVTEARDRAGGNIT 120

Qy    112 TVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLT 171
          ||||||||||||||||||||||||||||||||||||||||:|||||||| |||||| |||
Db    121 TVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPDAPRFVLWKDKLRPVPGKLT 180

Qy    172 DLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYV 231
          ||||||||||||||||||| |||||||||:||||||||||||| |||||||||||||||
```

```
Db    181 DLPFFDLMSIPGKLRAGFGAIGLRPSPPGYEESVEQFVRRNLGAEVFERLIEPFCSGVYA 240

Qy    232 GDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFR 291
          |||||| ||||||||||||:|||||||||||||||||| || |||||| ||||||||||
Db    241 GDPSKLIMKAAFGKVWKLEQTGGSIIGGTFKAIKERSSNPKPPRDPRLPTPKGQTVGSFR 300

Qy    292 KGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSY 351
          ||||||||||  |||||:||||||||||||||||| ||||||||||||:|||||||||||
Db    301 KGLRMLPDAICERLGSKVKLSWKLSSITKSEKGGYLLTYETPEGVVSLRSRSIVMTVPSY 360

Qy    352 VASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGV 411
          ||||||||||||||||||:|||||| ||||||||||||||||||||||||||||||:|||
Db    361 VASNILRPLSVAAADALSSFYYPPVAAVTISYPQEAIRDERLVDGELKGFGQLHPRSQGV 420

Qy    412 ETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKA 471
          ||||||||||||||||| |||||||||||| | ||:||||||||| ||||||||||||||
Db    421 ETLGTIYSSSLFPNRAPNGRVLLLNYIGGATNTEIVSKTESQLVEAVDRDLRKMLIKPKA 480

Qy    472 QDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGA 531
          ||| | |||||||||||||||||||| ||| |::||||:|||||||||||||||||||||
Db    481 QDPFVTGVRVWPQAIPQFLVGHLDTLGTAKTALSDNGLDGLFLGGNYVSGVALGRCVEGA 540

Qy    532 YEVASEVTGFLSRYAYK 548
          ||:|||||||||:||||
Db    541 YEIASEVTGFLSQYAYK 557


RESULT   4
Q9ZSI5
ID   Q9ZSI5      PRELIMINARY;      PRT;   545 AA.
AC   Q9ZSI5;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   T15B16.13 PROTEIN.
GN   T15B16.13.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   WASHU;
RT   "The A. thaliana Genome Sequencing Project.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Stoneking T., Smith R.;
RT   "The sequence of A. thaliana T15B16.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Waterston R.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF104919; AAC72870.1; -.
DR   Mendel; 39261; Arath;3076;39261.
DR   InterPro; IPR000209; -.
DR   InterPro; IPR001613; -.
DR   PRINTS; PR00757; AMINEOXDASEF.
DR   PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
SQ   SEQUENCE   545 AA;  58598 MW;  6AB825691A01BE70 CRC64;

  Query Match             72.5%;  Score 2061;  DB 10;  Length 545;
  Best Local Similarity   74.3%;  Pred. No. 1.3e-143;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   18;  Gaps    5;

Qy     22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
          |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db      3 LSLLRPTTQSLLPSFSKPN-LRLNVYKPLRLRCSVAGGPTVGSSKIEGGGGTTITTDCVI 61

Qy     77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
          || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db     62 VGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITREENGFLWEEGPNSFQPSDP 121

Qy    133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
          |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db    122 MLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTDLPFFDLMSIGGKIRAGFGAL 181

Qy    193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSG--------VYVGDPSKLSMKAAFG 244
          |:|||||| |||||:||||||| |||||||||||||        || |||||||||||||
Db    182 GIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGRVIDKTSCVYAGDPSKLSMKAAFG 241

Qy    245 KVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISAR 304
          ||||||: ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||
Db    242 KVWKLEQNGGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTVGSFRKGLRMLPEAISAR 301

Qy    305 LGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAA 364
          ||||:||||||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :|
Db    302 LGSKVKLSWKLSGITKLESGGYNLTYETPDGLVSVQSKSVVMTVPSHVASGLLRPLSESA 361

Qy    365 ADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFP 424
          |:|||  ||||| ||:||||:|||| | |:||||||||||||||||||||||||||||||
Db    362 ANALSKLYYPPVAAVSISYPKEAIRTECLIDGELKGFGQLHPRTQGVETLGTIYSSSLFP 421

Qy    425 NRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQ 484
          |||| ||:||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||
Db    422 NRAPPGRILLLNYIGGSTNTGILSKSEGELVEAVDRDLRKMLIKPNSTDPLKLGVRVWPQ 481

Qy    485 AIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSR 544
          ||||||||| | | |||:::  :| ||||||||||:||||||||||||| | || |:||
Db    482 AIPQFLVGHFDILDTAKSSLTSSGYEGLFLGGNYVAGVALGRCVEGAYETAIEVNNFMSR 541

Qy    545 YAYK 548
          ||||
Db    542 YAYK 545


RESULT   5
Q9SPL6
ID   Q9SPL6      PRELIMINARY;      PRT;   555 AA.
AC   Q9SPL6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PROTOPORPHYRINOGEN IX OXIDASE (EC 1.3.3.4).
GN   PPX1.
OS   Cichorium intybus (Chicory).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae;
OC   euasterids II; Asterales; Asteraceae; Lactuceae; Cichorium.
OX   NCBI_TaxID=13427;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Adomat C., Boeger P.;
RT   "Cloning, Sequence, Expression and Characterization of
RT   Protoporphyrinogen IX Oxidase from Chicory.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF160961; AAF00194.1; -.
DR   InterPro; IPR000209; -.
DR   PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
KW   Oxidoreductase.
SQ   SEQUENCE   555 AA;  60244 MW;  0AB31FA9DEED303D CRC64;

  Query Match             70.5%;  Score 2004.5;  DB 10;  Length 555;
  Best Local Similarity   71.5%;  Pred. No. 1.9e-139;
  Matches  383;  Conservative   60;  Mismatches   84;  Indels    9;  Gaps    3;

Qy     18 SSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTV-PSSAVDGGPAAELDCVI 76
          |   |   |    | :|   ::    || || |||:|||  : |  : :      |||||
Db     24 SGGSLTSKNPRYLITYSPAHRK----CNRWRFRCSIAKDSPITPPISNEFNSQPLLDCVI 79
```

```
Qy         77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
              |||||||||||| ::  :    |:::||||||| ||||:||||||||||||||||||||
Db         80 VGAGISGLCIAQALATKHASVSPDVIVTEARDRVGGNISTVERDGYLWEEGPNSFQPSDA 139

Qy        133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
              |||| || ||||||||||| ||||||| | |:|||||  ||||||||| ||||||||| :
Db        140 MLTMVVDSGLKDDLVLGDPTAPRFVLWGGDLKPVPSKPADLPFFDLMSFPGKLRAGFGAL 199

Qy        193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
              | |||||  |||||:||||||| ||||||||||||||| |||||||||||||||| ||:
Db        200 GFRPSPPDREESVEEFVRRNLGDEVFERLIEPFCSGVYAGDPSKLSMKAAFGKVWNLEQN 259

Qy        253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
              ||||:|| ||||::| :: | ||||||||||||||||||||  |||:||| ||||::||
Db        260 GGSIVGGAFKAIQDRKNSQKPPRDPRLPKPKGQTVGSFRKGQAMLPNAISTRLGSRVKLC 319

Qy        313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
              |||:||:| |  ||:||||||:|  |||:::|||||||||||::|||||: |||||| ||
Db        320 WKLTSISKLENRGYNLTYETPQGFESLQTKTIVMTVPSYVASDLLRPLSLGAADALSKFY 379

Qy        373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
              |||| ||:||||::||| :||:||:||||||||||:|||||||||||||||||||| |||
Db        380 YPPVAAVSISYPKDAIRADRLIDGQLKGFGQLHPRSQGVETLGTIYSSSLFPNRAPPGRV 439

Qy        433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
              ||||||||| ||||||||| ::|: |||||| |||:  |:||| :||||||:||||||:|
Db        440 LLLNYIGGATNPEILSKTEGEIVDAVDRDLRTMLIRRDAEDPLTLGVRVWPRAIPQFLIG 499

Qy        493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
              | | | :||||::  | :|:|||||||||||||:||| ||:||:||  |||:  ||
Db        500 HYDILDSAKAALSSGGFQGMFLGGNYVSGVALGKCVEAAYDVAAEVMNFLSQGVYK 555


RESULT   6
Q9M629
ID   Q9M629      PRELIMINARY;      PRT;   535 AA.
AC   Q9M629;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PROTOPORPHYRINOGEN IX OXIDASE.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC   Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. B73 INBRED;
RA   Volrath S.L.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF218052; AAF26417.1; -.
SQ   SEQUENCE   535 AA;  56835 MW;  9EEF8B0E2973F919 CRC64;

  Query Match             70.0%;  Score 1989.5;  DB 10;  Length 535;
  Best Local Similarity   72.3%;  Pred. No. 2.3e-138;
  Matches  394;  Conservative   49;  Mismatches   69;  Indels   33;  Gaps    6;

Qy         17 SSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVPSSAVDGGPA------- 69
              :::|||  || |  ||     :  || |                 :|| || |
Db         11 TAASPL--LNGTR-IPARLRHRGLSVRC-----------------AAVAGGAAEAPASTG 50

Qy         70 --AELDCVIVGAGISGLCIAQVMSANY--PNLMVTEARDRAGGNITTVER--DGYLWEEG 123
                   |||:|| |||||| || ::  :   :::||||| | |||||||||  :|||||||
Db         51 ARLSADCVVVGGGISGLCTAQALATRHGVGDVLVTEARARPGGNITTVERPEEGYLWEEG 110

Qy        124 PNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPG 183
              |||||||||:|||||| ||||||| |||||||||||:|||||||||  ||||||||||||
Db        111 PNSFQPSDPVLTMAVDSGLKDDLVFGDPNAPRFVLWEGKLRPVPSKPADLPFFDLMSIPG 170

Qy        184 KLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAF 243
              ||||| | :|:|| ||| |||||:||||||| ||||||||||||||| ||||||||||||
Db        171 KLRAGLGALGIRPPPPGREESVEEFVRRNLGAEVFERLIEPFCSGVYAGDPSKLSMKAAF 230

Qy        244 GKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISA 303
              ||||:|||||||||||| | |:|||  || ||| |||||||||| |||||| |||:||::
Db        231 GKVWRLEETGGSIIGGTIKTIQERSKNPKPPRDARLPKPKGQTVASFRKGLAMLPNAITS 290

Qy        304 RLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVA 363
               ||||:||||||:|||||:  || | |||||||||:|::|::||:|||||||||||||
Db        291 SLGSKVKLSWKLTSITKSDDKGYVLEYETPEGVVSVQAKSVIMTIPSYVASNILRPLSSD 350

Qy        364 AADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLF 423
              |||||| |||||| |||:|||:|||| | |:||||:||||||||:|||||||||||||||
Db        351 AADALSRFYYPPVAAVTVSYPKEAIRKECLIDGELQGFGQLHPRSQGVETLGTIYSSSLF 410

Qy        424 PNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWP 483
              ||||| |||||||||||| |  |:|||||:||| ||||||||||   | ||||:||||||
Db        411 PNRAPDGRVLLLNYIGGATNTGIVSKTESELVEAVDRDLRKMLINSTAVDPLVLGVRVWP 470

Qy        484 QAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLS 543
              |||||||||||| |  ||||::  | :|||||||||:||||||||||||| ||::: ||:
Db        471 QAIPQFLVGHLDLLEAAKAALDRGGYDGLFLGGNYVAGVALGRCVEGAYESASQISDFLT 530

Qy        544 RYAYK 548
              :||||
Db        531 KYAYK 535


RESULT   7
Q9LRI8
ID   Q9LRI8      PRELIMINARY;      PRT;   562 AA.
AC   Q9LRI8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE (PROTOX-I).
GN   SO-POX1.
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC   Caryophyllales; Chenopodiaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. CV. TONIC; TISSUE=LEAVES;
RA   Che F.S., Watanabe N., Iwano M., Inokuchi H., Takayama S., Yoshida S.,
RA   Isogai A.;
RT   "Molecular characterization and subcellular localization of
RT   protoporphyrinogen oxidase in spinach chloroplasts.";
RL   Plant Physiol. 0:0-0(2000).
DR   EMBL; AB029492; BAA96808.1; -.
DR   InterPro; IPR001613; -.
DR   InterPro; IPR001848; -.
DR   PRINTS; PR00757; AMINEOXDASEF.
DR   ProDom; PD001272; -; 1.
SQ   SEQUENCE   562 AA;  59929 MW;  638DCF6E63D259C4 CRC64;

  Query Match             69.1%;  Score 1966.5;  DB 10;  Length 562;
  Best Local Similarity   71.1%;  Pred. No. 1.2e-136;
  Matches  381;  Conservative   58;  Mismatches   84;  Indels   13;  Gaps    5;

Qy         26 NR-TSFIPFSSISKR--NSVNCNGWRTRCSVA----KDYTVPSSAVD--GGPAAELDCVI 76
              || |  || ||: :|  :|: |:   |  | |    ::  : :: ||  ||    |||||
Db         27 NRITILIPSSSLRRRGGSSIRCSTISTSNSAAAANYQNKNIGTNGVDGGGGGGGVLDCVI 86

Qy         77 VGAGISGLCIAQVMSANYPNL----MVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
              || ||||||||| :|  | ||    :||||:|| ||||||:| ||||||||||||||||
Db         87 VGGGISGLCIAQALSTKYSNLSTNFIVTEAKDRVGGNITTMEADGYLWEEGPNSFQPSDA 146

Qy        133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
              :|||||| |||::|||||||:|||||| |||||||||||||||||||| |||:||| | :
Db        147 VLTMAVDSGLKEELVLGDPNSPRFVLWNGKLRPVPSKLTDLPFFDLMSFPGKIRAGLGAL 206
```

```
Qy      193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
            ||||||| |||||||||||||| ||||||||||||||| |||||||||||||:|| ||:
Db      207 GLRPSPPAHEESVEQFVRRNLGDEVFERLIEPFCSGVYAGDPSKLSMKAAFGRVWVLEQK 266

Qy      253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
            |||||||| | |:||   || ||||||||||||||||||||| ||| ||| |||:|:|:|
Db      267 GGSIIGGTLKTIQERKDNPKPPRDPRLPKPKGQTVGSFRKGLSMLPTAISERLGNKVKVS 326

Qy      313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
            | || | ||  | |:||||||:|:||::::|:||||||||||::|||||  ||::|| |:
Db      327 WTLSGIAKSSNGEYNLTYETPDGLVSVRTKSVVMTVPSYVASSLLRPLSDVAAESLSKFH 386

Qy      373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
            |||| ||::|||:|||| | |:||||||||||| |:|||||||||||||||| ||| ||
Db      387 YPPVAAVSLSYPKEAIRSECLIDGELKGFGQLHSRSQGVETLGTIYSSSLFPGRAPPGRT 446

Qy      433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
            |:||||||  || || ||: :| | ||||||::|| | |: | |:|||||||||||||:|
Db      447 LILNYIGGDTNPGILDKTKDELAEAVDRDLRRILINPNAKAPRVLGVRVWPQAIPQFLIG 506

Qy      493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
            | | |  ||||: | | :|||||||||||||||||:||||| |:||  |||:|: |
Db      507 HFDLLDAAKAALTDGGHKGLFLGGNYVSGVALGRCIEGAYESAAEVVDFLSQYSDK 562


RESULT   8
Q9ZTA7
ID   Q9ZTA7       PRELIMINARY;      PRT;   563 AA.
AC   Q9ZTA7;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE PRECURSOR.
GN   PPX1.
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CC-621;
RX   MEDLINE=99077317; PubMed=9862501;
RA   Randolph-Anderson B.L., Sato R., Johnson A.M., Harris E.H.,
RA   Hauser C.R., Oeda K., Ishige F., Nishio S., Gillham N.W.,
RA   Boynton J.E.;
RT   "Isolation and characterization of a mutant protoporphyrinogen oxidase
RT   gene from Chlamydomonas reinhardtii conferring resistance to porphyric
RT   herbicides.";
RL   Plant Mol. Biol. 38:839-858(1998).
DR   EMBL; AF068635; AAC79685.1; -.
DR   Mendel; 39099; Chlre;3076;39099.
SQ   SEQUENCE   563 AA;  59802 MW;  FC5E8FFC934CAFC2 CRC64;

  Query Match             53.8%;  Score 1530.5;  DB 10;  Length 563;
  Best Local Similarity   59.5%;  Pred. No. 1.7e-104;
  Matches  300;  Conservative   73;  Mismatches  114;  Indels   17;  Gaps    6;

Qy       61 SSAVDGGPAAE-----------LDCVIVGAGISGLCIAQVMSANY--PNLMVTEARDRAG 107
            ::|  | | |             | ::|| |:|||   | ::| :   | :|||||:| |
Db       60 AAAATGAPTASGAGVAKTLDNVYDVIVVGGGLSGLVTGQALAAQHKIQNFLVTEARERVG 119

Qy      108 GNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVP 167
            ||||::  |||:||||||||||:| || :||| | : ||| ||| ||||| |:|||||||
Db      120 GNITSMSGDGYVWEEGPNSFQPNDSMLQIAVDSGCEKDLVFGDPTAPRFVWWEGKLRPVP 179

Qy      168 SKLTDLPFFDLMSIPGKLRAGFGPIGL-RPSPPGHEESVEQFVRRNLGGEVFERLIEPFC 226
            | | |   |||||||||:||| | |||   : |  |||||||:||||| ||| |||||||
Db      180 SGL-DAFTFDLMSIPGKIRAGLGAIGLINGAMPSFEESVEQFIRRNLGDEVFFRLIEPFC 238

Qy      227 SGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRL-PKPKGQ 285
            |||| |||||||||||| ::| ||: |||::||  |  :|| | |  |||||| ||||||
Db      239 SGVYAGDPSKLSMKAAFNRIWILEKNGGSLVGGAIKLFQERQSNPAPPRDPRLPPKPKGQ 298

Qy      286 TVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIV 345
            |||||||||:||||||   :  |::::||| |: :   | | | |:|||| | : :|::
Db      299 TVGSFRKGLKMLPDAIERNIPDKIRVNWKLVSLGREADGRYGLVYDTPEGRVKVFARAVA 358

Qy      346 MTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLV-DGELKGFGQL 404
            :| |||| :::::  : |||:|| :| ||||||||:|||  |:|:||   || : |||||
Db      359 LTAPSYVVADLVKEQAPAAAEALGSFDYPPVGAVTLSYPLSAVREERKASDGSVPGFGQL 418

Qy      405 HPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRK 464
            ||||||: |||||||||||| |||:| :||||||||  |  |:::|  |||| ||:|||
Db      419 HPRTQGITTLGTIYSSSLFPGRAPEGHMLLLNYIGGTTNRGIVNQTTEQLVEQVDKDLRN 478

Qy      465 MLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVAL 524
            |:||| |  | ||||||||:||||| :|||: |  |: |::  ||:|: |||||||||||
Db      479 MVIKPDAPKPRVVGVRVWPRAIPQFNLGHLEQLDKARKALDAAGLQGVHLGGNYVSGVAL 538

Qy      525 GRCVEGAYEVASEVTGFLSRYAYK 548
            |: ||  || |: :   :|: | |
Db      539 GKVVEHGYESAANLAKSVSKAAVK 562


RESULT   9
Q9T0P9
ID   Q9T0P9       PRELIMINARY;      PRT;   563 AA.
AC   Q9T0P9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HERBICIDE-RESISTANT PROTOPORPHYRINOGEN OXIDASE PRECURSOR.
GN   PPX1.
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GB-2674;
RX   MEDLINE=99077317; PubMed=9862501;
RA   Randolph-Anderson B.L., Sato R., Johnson A.M., Harris E.H.,
RA   Hauser C.R., Oeda K., Ishige F., Nishio S., Gillham N.W.,
RA   Boynton J.E.;
RT   "Isolation and Characterization of a Mutant Protoporphyrinogen Oxidase
RT   Gene from Chlamydomonas reinhardtii Conferring Resistance to Porphyric
RT   Herbicides.";
RL   Plant Mol. Biol. 38:839-858(1998).
DR   EMBL; AF030179; AAC79630.1; -.
KW   Transit peptide.
FT   TRANSIT       1     82       POTENTIAL.
FT   CHAIN        83    563       HERBICIDE-RESISTANT PROTOPORPHYRINOGEN
FT                                OXIDASE.
FT   VARIANT     389    389       M -> V.
SQ   SEQUENCE   563 AA;  59834 MW;  074584EC935CA3AF CRC64;

  Query Match             53.7%;  Score 1527.5;  DB 10;  Length 563;
  Best Local Similarity   59.3%;  Pred. No. 2.8e-104;
  Matches  299;  Conservative   74;  Mismatches  114;  Indels   17;  Gaps    6;

Qy       61 SSAVDGGPAAE-----------LDCVIVGAGISGLCIAQVMSANY--PNLMVTEARDRAG 107
            ::|  | | |             | ::|| |:|||   | ::| :   | :|||||:| |
Db       60 AAAATGAPTASGAGVAKTLDNVYDVIVVGGGLSGLVTGQALAAQHKIQNFLVTEARERVG 119

Qy      108 GNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVP 167
            ||||::  |||:||||||||||:| || :||| | : ||| ||| ||||| |:|||||||
Db      120 GNITSMSGDGYVWEEGPNSFQPNDSMLQIAVDSGCEKDLVFGDPTAPRFVWWEGKLRPVP 179

Qy      168 SKLTDLPFFDLMSIPGKLRAGFGPIGL-RPSPPGHEESVEQFVRRNLGGEVFERLIEPFC 226
            | | |   |||||||||:||| | |||   : |  |||||||:||||| ||| |||||||
Db      180 SGL-DAFTFDLMSIPGKIRAGLGAIGLINGAMPSFEESVEQFIRRNLGDEVFFRLIEPFC 238

Qy      227 SGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRL-PKPKGQ 285
```

```
            |||| |||||||||||| ::| ||: |||::||  |  :|| | |  |||||| ||||||
Db      239 SGVYAGDPSKLSMKAAFNRIWILEKNGGSLVGGAIKLFQERQSNPAPPRDPRLPPKPKGQ 298

Qy      286 TVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIV 345
            |||||||||:||||||   :  |::::||| |: :   | | | |:|||| | : :|::
Db      299 TVGSFRKGLKMLPDAIERNIPDKIRVNWKLVSLGREADGRYGLVYDTPEGRVKVFARAVA 358

Qy      346 MTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLV-DGELKGFGQL 404
            :| |||| :::::  : |||:|| :| |||:||||:|||  |:|:||   ||  : |||||
Db      359 LTAPSYVVADLVKEQAPAAAEALGSFDYPPMGAVTLSYPLSAVREERKASDGSVPGFGQL 418

Qy      405 HPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRK 464
            ||||||: |||||||||||| |||:| :|||||||||  |  |:::|  |||| ||:|||
Db      419 HPRTQGITTLGTIYSSSLFPGRAPEGHMLLLNYIGGTTNRGIVNQTTEQLVEQVDKDLRN 478

Qy      465 MLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVAL 524
            |:||| |  | ||||||||:||||| :|||: |  |: |::  ||:|: |||||||||||
Db      479 MVIKPDAPKPRVVGVRVWPRAIPQFNLGHLEQLDKARKALDAAGLQGVHLGGNYVSGVAL 538

Qy      525 GRCVEGAYEVASEVTGFLSRYAYK 548
            |: ||  || |: :   :|: | |
Db      539 GKVVEHGYESAANLAKSVSKAAVK 562
```

RESULT 10
Q9SLW5
ID   Q9SLW5   PRELIMINARY;   PRT;   502 AA.
AC   Q9SLW5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PROTOPORPHYRINOGEN IX OXIDASE.
GN   HEMG.
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20063348; PubMed=10594126;
RA   Kanjo N., Nishio S., Narita S., Oeda K., Inokuchi H.;
RT   "Nucleotide sequence of a cDNA clone encoding protoporphyrinogen IX
RT   oxidase (Accession No. AB025102) from soybean (PGR99-185).";
RL   Plant Physiol. 121:1383-1383(1999).
DR   EMBL; AB025102; BAA76348.1; -.
DR   InterPro; IPR000205; -.
DR   InterPro; IPR000759; -.
DR   InterPro; IPR002937; -.
DR   Pfam; PF01593; Amino_oxidase; 1.
DR   PRINTS; PR00419; ADXRDTASE.
SQ   SEQUENCE   502 AA; 55046 MW; EFC8DEF6E4DC2444 CRC64;

```
  Query Match             18.4%;  Score 522.5;  DB 10;  Length 502;
  Best Local Similarity   29.1%;  Pred. No. 2.7e-30;
  Matches  148;  Conservative   92;  Mismatches  225;  Indels   43;  Gaps   13;

Qy       61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120
            ||| |  | :     :||||:|||  |  : ::  ::  | ||  |||| : :| :|| :|
Db        3 SSATDDNPRSVKRVAVVGAGVSGLAAAYKLKSHGLDVTVFEAEGRAGGRLRSVSQDGLIW 62

Qy      121 EEGPNSFQPSDPMLTMAVDC-GLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLM 179
            :|| |:   |:  :   :|  ||::          |:::  |    ||:    |    |:
Db       63 DEGANTMTESEIEVKGLIDALGLQEKQQFPISQHKRYIVKNGAPLLVPTNPAALLKSKLL 122

Qy      180 SIPGKLRAGFGPIGLRPSPPGH------EESVEQFVRRNLGGEVFERLIEPFCSGVYVGD 233
            |   |:   | |   : | | :       ||| :|  |: | || : ||:||  |    |
Db      123 SAQSKIHLIFEPFMWKRSDPSNVCDENSVESVGRFFERHFGKEVVDYLIDPFVGGTSAAD 182

Qy      234 PSKLSMKAAFGKVWKLEETGGSIIGGTF--KAIKERSSTPKAPRDPRLPKPKGQTVGSFR 291
            |  |||: :| ::| ||:  |||| |    |   :|  | :     |  | |  :  ||:
Db      183 PESLSMRHSFPELWNLEKRFGSIIAGALQSKLFAKREKTGENRTALRKNKHKRGSF-SFQ 241
```

```
Qy      292 KGLRMLPDAISARLG-SKLKLSWKLSSITKSEKG-----GYHLTYETPEGVVSLQSRSIV 345
             |:: | | :   ||    |||: |: ::     |      : :|  : :    :   :::
Db      242 GGMQTLTDTLCKELGKDDLKLNEKVLTLAYGHDGSSSSQNWSITSASNQSTQDVD--AVI 299

Qy      346 MTVPSYVASNIL-------RPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGEL 398
            || | |   :|         ||:      |    | |:  :  ::  :|  ::        |
Db      300 MTAPLYNVKDIKITKRGTPFPLNF-----LPEVSYVPISVMITTFKKENVK------RPL 348

Qy      399 KGFGQLHP---RTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLV 455
            :||| | |   :  |::||||::|| :||:|||    |   :|||  :| |:   :  :|
Db      349 EGFGVLVPSKEQKNGLKTLGTLFSSMMFPDRAPSDLYLYTTFIGGTQNRELAQASTDELR 408

Qy      456 EVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLG 515
            ::|  ||||:|      :|  |    | :  | :   :   |        |  | | |
Db      409 KIVTSDLRKLL--GAEGEPTFVNHFYWSKGFPLYGRNYGSVLQAIDKIEKD--LPGFFFA 464

Qy      516 GNYVSGVALGRCVEGAYEVASEVTGFLS 543
            |||  |:::|: :    : |  |  :|:
Db      465 GNYKGGLSVGKAIASGCKAADLVISYLN 492
```

RESULT 11
O67814
ID   O67814   PRELIMINARY;   PRT;   436 AA.
AC   O67814;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE.
GN   HEMG.
OS   Aquifex aeolicus.
OC   Bacteria; Aquificales; Aquificaceae; Aquifex.
OX   NCBI_TaxID=63363;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VF5;
RX   MEDLINE=98196666; PubMed=9537320;
RA   Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA   Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA   Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT   "The complete genome of the hyperthermophilic bacterium Aquifex
RT   aeolicus.";
RL   Nature 392:353-358(1998).
DR   EMBL; AE000768; AAC07778.1; -.
DR   InterPro; IPR000205; -.
DR   InterPro; IPR002937; -.
DR   Pfam; PF01593; Amino_oxidase; 1.
SQ   SEQUENCE   436 AA; 48987 MW; ED5F2B1BCD1DEEF7 CRC64;

```
  Query Match             18.1%;  Score 515;  DB 2;  Length 436;
  Best Local Similarity   29.9%;  Pred. No. 7.7e-30;
  Matches  144;  Conservative  103;  Mismatches  170;  Indels   64;  Gaps   16;

Qy       73 DCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
            : |::|||||||  |  :     :::| |  || || | ||:  |||:| |  :
Db        5 EVVVIGAGISGLSTAYRLKKEGVDVVVYEKDDRIGGTIHTVKEKGYLFEVGAQTILADQE 64

Qy      133 MLTMAVDCGLKDDLVLGDPNAP-RFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGP 191
            ::    : |::   |   |::  |::  ||:| |:|    :     |:|:  ||:
Db       65 VIDFLKEAGIEP--VEASPSSKYRYIYKKGRLIPLPMSPVEFLKTPLLSLKTKLKV-LTE 121

Qy      192 IGLRPSPPGHEE--SVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKL 249
            |  |    | :|  |:  ||| : | |    :: || |||| ||| |||:| |  |:::
Db      122 IFKR----GVDEDISIADFVREHFGEEFLNYVVAPFISGVYAGDPEKLSLKHATPKLYEA 177

Qy      250 EETGGSIIGGTFKA-IKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARL--- 305
            ::  ||:|    || |||:::            |||: : || :||  | :|:: :|
Db      178 QKKYGSLI----KAFIKEKTA-----------GPKGKLI-SFGEGLGELINALAQKLEVH 221

Qy      306 --GSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVA 363
                 |::  |     : :  |             ::::|:|:  |:| :| :|: :| :
```

```
Db      222 TENVVLRMR-KFEDFFRLDVRG-----------KKVETKSVVVASPAYTSSYLLKEVSFS 269

Qy      364 AADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLF 423
            |::      ||||  | :            |:|:      |   :     || :: | ||
Db      270 ASEEFDKIDYPPVVVVNVG-----------VEGKFPKVRLLSSESGKKRILGAMFMSKLF 318

Qy      424 PNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLV--VGVRV 481
            | |||:|: ||  ::||| : |::  :| ::  :|:|:|:::|     |   :  : |:
Db      319 PGRAPQGKELLTVFLGGATDREVIELSEEEIENIVERELKEIL-----QIDCIDFMHVQK 373

Qy      482 WPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGF 541
            | :||||: :|:   |: |:    |    |||| ||:: ||:   |:  : :|| :|  |
Db      374 WKRAIPQYTLGYDRFLNLAQEMEKD--YPGLFLTGNWLYGVSTADCIRASKKVAQKVLSF 431

Qy      542 L 542
            |
Db      432 L 432


RESULT  12
Q9FYV8
ID   Q9FYV8      PRELIMINARY;      PRT;   544 AA.
AC   Q9FYV8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PROTOPORPHYRINOGEN IX OXIDASE.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC   Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Volrath S.L.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF273767; AAG00946.1; -.
SQ   SEQUENCE   544 AA;  59020 MW;  101CEF56EB9A151B CRC64;

  Query Match             16.1%;  Score 459;  DB 10;  Length 544;
  Best Local Similarity   28.3%;  Pred. No. 1.5e-25;
  Matches  142;  Conservative   88;  Mismatches  226;  Indels   46;  Gaps   16;

Qy       76 IVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDPMLT 135
            :||||:|||  |  :  :  |: | || ||||| | |    |::|:|| |:    :   :
Db       50 VVGAGVSGLAAAYRLRQSGVNVTVFEAADRAGGKIRTNSEGGFVWDEGANTMTEGEWEAS 109

Qy      136 MAV-DCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI-- 192
              : | ||:|     :    |:::  |    :||    |    ::|   |:   | |
Db      110 RLIDDLGLQDKQQYPNSQHKRYIVKDGAPALIPSDPISLMKSSVLSTKSKIALFFEPFLY 169

Qy      193 ---GLRPSPPGHE----ESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGK 245
                 | |    |    |||  |  |: | || :  ::|| :|   |||  ||:: ||
Db      170 KKANTRNSGKVSEEHLSESVGSFCERHFGREVVDYFVDPFVAGTSAGDPESLSIRHAFPA 229

Qy      246 VWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVG-SFRKGLRMLPDAISAR 304
            :| ||   ||:| |   :       |   |     | : : |  ||  |:: | :|:
Db      230 LWNLERKYGSVIVGAILSKLAAKGDPVKTRHDSSGKRRNRRVSFSFHGGMQSLINALHNE 289

Qy      305 LG-SKLKLSWKLSSITKS-----EKGGYHLTYETPE-GVVSLQSR----SIVMTVPSYVA 353
            :|   :||  :: |:  :       | : :: :: : |   | |     :::|| |
Db      290 VGDDNVKLGTEVLSLACTFDGVPALGRWSISVDSKDSGDKDLASNQTFDAVIMTAP---L 346

Qy      354 SNILRPL-----SVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHP-- 406
            ||: |       :    | |    | |:  :  :: :: ::        |:||| | |
Db      347 SNVRRMKFTKGGAPVVLDFLPKMDYLPLSLMVTAFKKDDVKK------PLEGFGVLIPYK 400

Qy      407 --RTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRK 464
              :  |::||||::|| :||:|||  : |   ::||: | ::     | | ::|  ||:|
Db      401 EQQKHGLKTLGTLFSSMMFPDRAPDDQYLYTTFVGGSHNRDLAGAPTSILKQLVTSDLKK 460

Qy      465 MLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMN-DNGLEGLFLGGNYVSGVA 523
```

```
            :|       |  |    |  | |  | || |  |  :|    :  | | |  ||   |:|
Db      461 LL--GVEGQPTFVKHVYWGNAFP--LYGH-DYSSVLEAIEKMEKNLPGFFYAGNSKDGLA 515

Qy      524 LGRCVEGAYEVASEVTGFLSRY 545
            :|  :    : |     :|  :
Db      516 VGSVIASGSKAADLAISYLESH 537


RESULT  13
Q9LYA7
ID   Q9LYA7      PRELIMINARY;      PRT;   501 AA.
AC   Q9LYA7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE-LIKE PROTEIN.
GN   F18O22_10.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA   Rudd S., Lemcke K., Mayer K.F.X.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL163817; CAB87761.1; -.
DR   InterPro; IPR000205; -.
DR   InterPro; IPR000759; -.
DR   PRINTS; PR00419; ADXRDTASE.
SQ   SEQUENCE   501 AA;  54829 MW;  50CB8E1F3C0ED2F0 CRC64;

  Query Match             16.0%;  Score 456;  DB 10;  Length 501;
  Best Local Similarity   27.0%;  Pred. No. 2.1e-25;
  Matches  142;  Conservative   95;  Mismatches  194;  Indels   94;  Gaps   18;

Qy       63 AVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLWEE 122
            || |   |     :||||:|||  |  : :   |: | ||  | || : :| ::| :|:|
Db       13 AVSGKRVA-----VVGAGVSGLAAAYKLKSRGLNVTVFEADGRVGGKLRSVMQNGLIWDE 67

Qy      123 GPNSFQPSDPML-TMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSI 181
            | |:   ::| : ::  | ||::          |:::  |    :|:   :|    ::|
Db       68 GANTMTEAEPEVGSLLDDLGLREKQQFPISQKKRYIVRNGVPVMLPTNPIELVTSSVLST 127

Qy      182 PGKLRAGFGPI-----GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSK 236
              | :    |        : |    |||| :| :|: | || : ||:||  |    ||
Db      128 QSKFQILLEPFLWKKKSSKVSDASAEESVSEFFQRHFGQEVVDYLIDPFVGGTSAADPDS 187

Qy      237 LSMKAAFGKVWKLEETGGSII-----------GGTFKAIKERSSTPKAPRDPRLPKPKGQ 285
            |||| :|  :|    : ||||           ||  :  |    | |  |
Db      188 LSMKHSFPDLW---NSFGSIIVGAIRTKFAAKGGKSRDTKSSPGTKKGSRGS-------- 236

Qy      286 TVGSFRKGLRMLPDAI-------SARLGSK-LKLS---------WKLSSITKSEKGGYHL 328
               ||: |:::||| :          | || | ||         | || :: :|
Db      237 --FSFKGGMQILPDTLCKSLSHDEINLDSKVLSLSYNSGSRQENWSLSCVSHNE------ 288

Qy      329 TYETPEGVVSLQSRSIVMTVPSYVASNI--LRPLSVAAA------DALSNFYYPPVGAVT 380
                       | ::     | | |: :  :: : |         : |    | |:  :
Db      289 ----------TQRQN-----PHYDAAPLCNVKEMKVMKGGQPFQLNFLPEINYMPLSVLI 333

Qy      381 ISYPQEAIRDERLVDGELKGFGQLHP---RTQGVETLGTIYSSSLFPNRAPKGRVLLLNY 437
             :: :| ::        |:||| | |   :  | :||||::|| :||:|:|    |   :
Db      334 TTFTKEKVK------RPLEGFGVLIPSKEQKHGFKTLGTLFSSMMFPDRSPSDVHLYTTF 387

Qy      438 IGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTL 497
            |||::| |:   :  :| :||  ||:::|      :|: |    | :| | :   : |::
Db      388 IGGSRNQELAKASTDELKQVVTSDLQRLL--GVEGEPVSVNHYYWRKAFPLYDSSY-DSV 444
```

```
Qy    498 STAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFL 542
            |   | :| | | |  ||:  |:::|: :    : |  |  :|
Db    445 MEAIDKM-ENDLPGFFYAGNHRGGLSVGKSIASGCKAADLVISYL 488


RESULT  14
Q9FMS9
ID   Q9FMS9      PRELIMINARY;      PRT;   547 AA.
AC   Q9FMS9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PROTOPORPHYRINOGEN IX OXIDASE.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX.  NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   MEDLINE=98162728; PubMed=9501997;
RA   Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT   Sequence features of the regions of 1,191,918 bp covered by seventeen
RT   physically assigned P1 clones.";
RL   DNA Res. 4:401-414(1997).
DR   EMBL; AB007650; BAB08301.1; -.
SQ   SEQUENCE   547 AA;  60266 MW;  5474575A8B707A6E CRC64;

  Query Match             16.0%;  Score 456;  DB 10;  Length 547;
  Best Local Similarity   27.0%;  Pred. No. 2.5e-25;
  Matches  142;  Conservative   95;  Mismatches  194;  Indels   94;  Gaps   18;

Qy     63 AVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLWEE 122
          || |   |     :||||:|||  |  : :   |: | ||  | ||  : :|  ::| :|:|
Db     59 AVSGKRVA-----VVGAGVSGLAAAYKLKSRGLNVTVFEADGRVGGKLRSVMQNGLIWDE 113

Qy    123 GPNSFQPSDPML-TMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSI 181
          | |:   ::| : ::  | ||::          |:::  |    :|:   :|    ::|
Db    114 GANTMTEAEPEVGSLLDDLGLREKQQFPISQKKRYIVRNGVPVMLPTNPIELVTSSVLST 173

Qy    182 PGKLRAGFGPI-----GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSK 236
            | :    |        : |    |||| :| :|: | || : ||:||  |    ||
Db    174 QSKFQILLEPFLWKKKSSKVSDASAEESVSEFFQRHFGQEVVDYLIDPFVGGTSAADPDS 233

Qy    237 LSMKAAFGKVWKLEETGGSII-----------GGTFKAIKERSSTPKAPRDPRLPKPKGQ 285
          |||| :|  :|    : ||||           ||  :  |    | |  |
Db    234 LSMKHSFPDLW---NSFGSIIVGAIRTKFAAKGGKSRDTKSSPGTKKGSRGS-------- 282

Qy    286 TVGSFRKGLRMLPDAI-------SARLGSK-LKLS---------WKLSSITKSEKGGYHL 328
             ||: |:::||| :          | || | ||         | ||  :: :|
Db    283 --FSFKGGMQILPDTLCKSLSHDEINLDSKVLSLSYNSGSRQENWSLSCVSHNE------ 334

Qy    329 TYETPEGVVSLQSRSIVMTVPSYVASNI--LRPLSVAAA------DALSNFYYPPVGAVT 380
                     | ::     | | |: :  :: : |         : |    | |:  :
Db    335 ----------TQRQN-----PHYDAAPLCNVKEMKVMKGGQPFQLNFLPEINYMPLSVLI 379

Qy    381 ISYPQEAIRDERLVDGELKGFGQLHP---RTQGVETLGTIYSSSLFPNRAPKGRVLLLNY 437
           :: :| ::        |:||| | |   :  | :||||::|| :||:|:|    |   :
Db    380 TTFTKEKVK------RPLEGFGVLIPSKEQKHGFKTLGTLFSSMMFPDRSPSDVHLYTTF 433

Qy    438 IGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTL 497
          |||::| |:   :  :| :||  ||:::|      :|: |    | :| | :   : |::
Db    434 IGGSRNQELAKASTDELKQVVTSDLQRLL--GVEGEPVSVNHYYWRKAFPLYDSSY-DSV 490

Qy    498 STAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFL 542
            |   | :| | | |  ||:  |:::|: :    : |  |  :|
Db    491 MEAIDKM-ENDLPGFFYAGNHRGGLSVGKSIASGCKAADLVISYL 534
```

```
RESULT  15
O64385
ID   O64385      PRELIMINARY;      PRT;   404 AA.
AC   O64385;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4).
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4113;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. BINTJE;
RA   Johnston D.J., Droz E., Rochaix J.D., Malnoe P.C.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC       IX + H(2)O.
DR   EMBL; AJ225108; CAA12401.1; -.
DR   Mendel; 29309; Soltu;3077;29309.
DR   InterPro; IPR000205; -.
KW   Oxidoreductase.
SQ   SEQUENCE   404 AA;  44454 MW;  A16C21D30AC46A34 CRC64;

  Query Match             14.9%;  Score 424;  DB 10;  Length 404;
  Best Local Similarity   28.7%;  Pred. No. 3.5e-23;
  Matches  123;  Conservative   83;  Mismatches  152;  Indels   70;  Gaps   15;

Qy     60 PSSAVDGGPAAELDC----VIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVER 115
          ||:  |     : :|     ::|||:|||  |  :  :  |: | ||  |||| : :: :
Db      3 PSAGED-----KQNCPKRVAVIGAGVSGLAAAYKLKIHGLNVTVFEAEGRAGGKLRSLSQ 57

Qy    116 DGYLWEEGPNSFQPSDPMLTMAVD-CGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLP 174
          || :|:|| |:   |:  :|  :|  ||::          |::   |    :||   ||
Db     58 DGLIWDEGANTMTESEGDVTFLLDSLGLREKQQFPLSQNKRYIARNGTPTLIPSNPIDLI 117

Qy    175 FFDLMSIPGKLRAGFGPIGLRPSP----PGHEESVEQFVRRNLGGEVFERLIEPFCSGVY 230
            : :|   ||:  | |:  : :         |||  | :|: | || : ||:|| :|
Db    118 KSNFLSTGSKLQMLFEPLLWKNNKLTKVSDEHESVSGFFQRHFGKEVVDYLIDPFVAGTC 177

Qy    231 VGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKA----IKERSSTPKAPRDPRLPKPKGQT 286
           |||  |||  :| ::| ||:  ||:| |  ::    |||:   |     |:    | :
Db    178 GGDPDSLSMHLSFPELWNLEKRFGSVIVGAIRSKLSPIKEKKQGP-----PKTSVNKKRQ 232

Qy    287 VGSFR--KGLRMLPDAI-------SARLGSK-LKL-----------SWKLSSITKSEKGG 325
           |||    |:: | |||         || |: |:|           || : | :  ::
Db    233 RGSFSFLGGMQTLTDAICKDLKEDELRLNSRVLELSCSCSGDSAIDSWSIFSASPHKRQA 292

Qy    326 YHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY----YPPVGAVTI 381
             :::           :::|| |     ::   ::      | ||     | |:  |
Db    293 EEESFD-----------AVIMTAPLCDVKSM--KIAKRGNPFLLNFIPEVDYVPLSVVIT 339

Qy    382 SYPQEAIRDERLVDGELKGFGQLHP---RTQGVETLGTIYSSSLFPNRAPKGRVLLLNYI 438
          :: :|:::        |:||| | |   :  |::||||::|| :||:|||    |   ::
Db    340 TFKKESVKH------PLEGFGVLVPSQEQKHGLKTLGTLFSSMMFPDRAPNNVYLYTTFV 393

Qy    439 GGAKNPEI 446
          ||::| |:
Db    394 GGSRNREL 401


Search completed: July  3, 2001, 10:42:33
Job time: 571 sec
```

```
                        GenCore version 4.5
              Copyright (c) 1993 - 2000  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          July  3, 2001, 10:27:57 ; Search time 61.92 Seconds
                                             (without alignments)
                                             674.155 Million cell updates/sec

Title:           US-09-508-418-2
Perfect score:   2844
Sequence:        1 MTTTPIANHPNIFTHQSSSS..........EGAYEVASEVTGFLSRYAYK 548

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        PIR_68:*
                 1:  pir1:*
                 2:  pir2:*
                 3:  pir3:*
                 4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                        %
Result               Query
  No.     Score     Match Length DB  ID                       Description
 ---------------------------------------------------------------------------
     1     2821     99.2    548   2  T04058                   protoporphyrinogen
     2   2565.5     90.2    557   2  T07116                   protoporphyrinogen
     3     2075     73.0    537   2  G85021                   protoporphyrinogen
     4     2061     72.5    545   2  T02005                   protoporphyrinogen
     5      515     18.1    436   2  B70473                   protoporphyrinogen
     6      505     17.8    504   2  T04076                   protoporphyrinogen
     7      485     17.1    470   2  D47045                   coproporphyrinogen
     8      456     16.0    501   2  T48595                   protoporphyrinogen
     9      424     14.9    404   2  T07121                   protoporphyrinogen
    10    407.5     14.3    467   2  D83800                   protoporphyrinogen
    11    368.5     13.0    477   2  S65684                   protoporphyrinogen
    12    363.5     12.8    462   2  A75435                   protoporphyrinogen
    13    363.5     12.8    477   2  A56449                   protoporphyrinogen
    14    362.5     12.7    477   2  S68367                   protoporphyrinogen
    15      317     11.1    468   2  T34679                   probable protoporp
    16    302.5     10.6    490   2  T38088                   protoporphyrinogen
    17      254      8.9    435   2  F86601                   protoporphyrinogen
    18      254      8.9    435   2  A81518                   protoporphyrinogen
    19      249      8.8    435   2  D72023                   protoporphyrinogen
    20    231.5      8.1    424   2  D81739                   protoporphyrinogen
    21      221      7.8    426   2  F70869                   probable hemY' - M
    22      211      7.4    424   2  A71476                   probable protoporp
    23      189      6.6    539   2  S50472                   hypothetical prote
    24    174.5      6.1    421   2  C75327                   oxidoreductase, FA
    25    169.5      6.0    441   2  A75361                   hypothetical prote
    26    153.5      5.4    416   2  T37023                   probable oxidoredu
    27      148      5.2    440   2  B71153                   hypothetical prote
    28      146      5.1    430   2  E84212                   hypothetical prote
    29      144      5.1    488   2  T47787                   hypothetical prote
```

```
    30      143      5.0    490   2  A84861                   probable amine oxi
    31      141      5.0    448   2  H70947                   hypothetical prote
    32      139      4.9    570   2  S54134                   phytoene dehydroge
    33    136.5      4.8    467   2  T35307                   probable oxidoredu
    34      134      4.7    499   2  I51346                   monoamine oxidase
    35    132.5      4.7    527   2  S03974                   amine oxidase (fla
    36      131      4.6    566   2  F71403                   hypothetical prote
    37      127      4.5    574   2  A96612                   hypothetical prote
    38    125.5      4.4    472   2  S74886                   phytoene dehydroge
    39    123.5      4.3    526   2  JT0528                   amine oxidase (fla
    40      123      4.3    582   2  S29314                   phytoene dehydroge
    41    122.5      4.3    489   2  A47259                   corticosteroid-bin
    42      122      4.3    471   2  S76290                   hypothetical prote
    43      121      4.3    404   2  A69467                   hypothetical prote
    44    119.5      4.2    472   2  T35936                   probable dehydroge
    45    119.5      4.2    478   2  A47693                   putrescine oxidase


                                 ALIGNMENTS


RESULT   1
T04058
protoporphyrinogen oxidase (EC 1.3.3.4) IX precursor, chloroplast - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C;Accession: T04058
R;Lermontova, I.; Kruse, E.; Mock, H.P.; Grimm, B.
Proc. Natl. Acad. Sci. U.S.A. 94, 8895-8900, 1997
A;Title: Cloning and characterization of a plastidal and a mitochondrial isoform of t
A;Reference number: Z15186; MUID:97385200
A;Accession: T04058
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-548 <LER>
A;Cross-references: EMBL:Y13465; NID:g2370332; PIDN:CAA73865.1; PID:g2370333
A;Experimental source: strain SR1
C;Genetics:
A;Gene: ppxI
C;Function:
A;Pathway: tetrapyrrole synthesis
C;Keywords: chloroplast; oxidoreductase
F;1-50/Domain: transit peptide (chloroplast) #status predicted <TNP>
F;51-548/Product: protoporphyrinogen oxidase IX #status predicted <MAT>


  Query Match             99.2%;  Score 2821;  DB 2;  Length 548;
  Best Local Similarity   99.3%;  Pred. No. 2.1e-200;
  Matches  544;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60

Qy         61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120

Qy        121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLAFFDLMS 180

Qy        181 IPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMK 240
              |||||||||| ||||||||||||||||||||||||||||||||||||||| |||||||||
Db        181 IPGKLRAGFGAIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYAGDPSKLSMK 240

Qy        241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300

Qy        301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360
```

```
Qy     361 SVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420
           |||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db     361 SVAAADALSNFYYPPVGAVTITYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420

Qy     421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480

Qy     481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540

Qy     541 FLSRYAYK 548
           ||||||||
Db     541 FLSRYAYK 548

RESULT   2
T07116
protoporphyrinogen oxidase (EC 1.3.3.4) IX, chloroplast - potato
C;Species: Solanum tuberosum (potato)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999
C;Accession: T07116
R;Johnston, D.J.
submitted to the EMBL Data Library, April 1998
A;Reference number: Z15932
A;Accession: T07116
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-557 <JOH>
A;Cross-references: EMBL:AJ225107; NID:e1289157; PIDN:CAA12400.1; PID:e1289158
A;Experimental source: cv. Bintje
C;Genetics:
A;Genome: nuclear
C;Function:
A;Pathway: tetrapyrrole biosynthesis
C;Keywords: chloroplast; oxidoreductase


  Query Match             90.2%;  Score 2565.5;  DB 2;  Length 557;
  Best Local Similarity   88.9%;  Pred. No. 1.5e-181;
  Matches  495;  Conservative   21;  Mismatches   32;  Indels    9;  Gaps    2;

Qy       1 MTTTPIANHPNIFTHQ--------SSSSP-LAFLNRTSFIPFSSISKRNSVNCNGWRTRC 51
           |||| :||||:||||:        |||||   |||||:|||: | |||||||||||||||
Db       1 MTTTAVANHPSIFTHRSPLPSPSSSSSSPSFLFLNRTNFIPYFSTSKRNSVNCNGWRTRC 60

Qy      52 SVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNIT 111
           ||||||||| | |||    |||||:|||||||||||:|:|||||||||||||||||||||
Db      61 SVAKDYTVPPSEVDGNQFPELDCVVVGAGISGLCIAKVISANYPNLMVTEARDRAGGNIT 120

Qy     112 TVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLT 171
           ||||||||||||||||||||||||||||||||||||||||:|||||||| |||||| |||
Db     121 TVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPDAPRFVLWKDKLRPVPGKLT 180

Qy     172 DLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYV 231
           ||||||||||||||||||| |||||||||:||||||||||||| ||||||||||||||| 
Db     181 DLPFFDLMSIPGKLRAGFGAIGLRPSPPGYEESVEQFVRRNLGAEVFERLIEPFCSGVYA 240

Qy     232 GDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFR 291
           |||||| ||||||||||||:|||||||||||||||||| || ||||||| ||||||||||
Db     241 GDPSKLIMKAAFGKVWKLEQTGGSIIGGTFKAIKERSSNPKPPRDPRLPTPKGQTVGSFR 300

Qy     292 KGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSY 351
           ||||||||||  |||||:||||||||||||||||| ||||||||||||:|||||||||||
Db     301 KGLRMLPDAICERLGSKVKLSWKLSSITKSEKGGYLLTYETPEGVVSLRSRSIVMTVPSY 360

Qy     352 VASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGV 411
           ||||||||||||||||||:|||||| ||||||||||||||||||||||||||||||:|||
Db     361 VASNILRPLSVAAADALSSFYYPPVAAVTISYPQEAIRDERLVDGELKGFGQLHPRSQGV 420

Qy     412 ETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKA 471
           ||||||||||||||||| |||||||||||| | ||:||||||||| ||||||||||||||
Db     421 ETLGTIYSSSLFPNRAPNGRVLLLNYIGGATNTEIVSKTESQLVEAVDRDLRKMLIKPKA 480

Qy     472 QDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGA 531
           ||| | |||||||||||||||||||| ||| |::||||:|||||||||||||||||||||
Db     481 QDPFVTGVRVWPQAIPQFLVGHLDTLGTAKTALSDNGLDGLFLGGNYVSGVALGRCVEGA 540

Qy     532 YEVASEVTGFLSRYAYK 548
           ||:|||||||||:||||
Db     541 YEIASEVTGFLSQYAYK 557

RESULT   3
G85021
protoporphyrinogen oxidase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: G85021
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488
A;Accession: G85021
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-537 <STO>
A;Cross-references: GB:NC_001268; NID:g7268212; PIDN:CAB77739.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g01690
A;Map position: 4


  Query Match             73.0%;  Score 2075;  DB 2;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 2.4e-145;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy      22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
           |: |  |:     | || | :  | :   | |||||   || || ::||     :  ||||
Db       3 LSLLRPTTQSLLPSFSKPN-LRLNVYKPLRLRCSVAGGPTVGSSKIEGGGGTTITTDCVI 61

Qy      77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
           || ||||||||| ::   :    |||:||||:|| |||| | | :|:||||||||||||||
Db      62 VGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITREENGFLWEEGPNSFQPSDP 121

Qy     133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
           |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db     122 MLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTDLPFFDLMSIGGKIRAGFGAL 181

Qy     193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
           |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||: 
Db     182 GIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGVYAGDPSKLSMKAAFGKVWKLEQN 241

Qy     253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
           ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db     242 GGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTVGSFRKGLRMLPEAISARLGSKVKLS 301

Qy     313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
           |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db     302 WKLSGITKLESGGYNLTYETPDGLVSVQSKSVVMTVPSHVASGLLRPLSESAANALSKLY 361

Qy     373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
           |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db     362 YPPVAAVSISYPKEAIRTECLIDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPPGRI 421

Qy     433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
           ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db     422 LLLNYIGGSTNTGILSKSEGELVEAVDRDLRKMLIKPNSTDPLKLGVRVWPQAIPQFLVG 481

Qy     493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
           | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db     482 HFDILDTAKSSLTSSGYEGLFLGGNYVAGVALGRCVEGAYETAIEVNNFMSRYAYK 537
```

```
RESULT   4
T02005
protoporphyrinogen oxidase (EC 1.3.3.4) - Arabidopsis thaliana
N;Alternate names: protein T15B16.13
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 21-Jul-2000
C;Accession: T02005; JC5488
R;Stoneking, T.; Smith, R.
submitted to the EMBL Data Library, November 1998
A;Description: The sequence of A. thaliana T15B16.
A;Reference number: Z14488
A;Accession: T02005
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-545 <STO>
A;Cross-references: EMBL:AF104919; NID:g3859590; PID:g3859604
A;Experimental source: cultivar Columbia
R;Narita, S.; Tanaka, R.; Ito, T.; Okada, K.; Taketani, S.; Inokuchi, H.
Gene 182, 169-175, 1996
A;Title: Molecular cloning and characterization of a cDNA that encodes protoporphyrinoge
A;Reference number: JC5488; MUID:97136707
A;Accession: JC5488
A;Molecule type: mRNA
A;Residues: 1-217,226-545 <NAR>
A;Cross-references: DDBJ:D83139; NID:g1183006; PIDN:BAA11820.1; PID:g1877018
C;Genetics:
A;Map position: 4
A;Introns: 126/3; 188/3; 225/1; 274/2; 301/2; 358/3; 412/1; 446/3
A;Note: T15B16.13
C;Function:
A;Description: catalyzes the removal of six electrons from protoporphyrinogen to generat
A;Pathway: chlorophyll biosynthetisis; heme biosynthetisis
C;Keywords: oxidoreductase


  Query Match             72.5%;  Score 2061;  DB 2;  Length 545;
  Best Local Similarity   74.3%;  Pred. No. 2.6e-144;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   18;  Gaps    5;

Qy      22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
           |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db       3 LSLLRPTTQSLLPSFSKPN-LRLNVYKPLRLRCSVAGGPTVGSSKIEGGGGTTITTDCVI 61

Qy      77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
           || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db      62 VGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITREENGFLWEEGPNSFQPSDP 121

Qy     133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
           |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db     122 MLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTDLPFFDLMSIGGKIRAGFGAL 181

Qy     193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSG--------VYVGDPSKLSMKAAFG 244
           |:|||||| |||||:||||||| |||||||||||||        || |||||||||||||
Db     182 GIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGRVIDKTSCVYAGDPSKLSMKAAFG 241

Qy     245 KVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISAR 304
           ||||||: ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||
Db     242 KVWKLEQNGGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTVGSFRKGLRMLPEAISAR 301

Qy     305 LGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAA 364
           ||||:||||||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :|
Db     302 LGSKVKLSWKLSGITKLESGGYNLTYETPDGLVSVQSKSVVMTVPSHVASGLLRPLSESA 361

Qy     365 ADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFP 424
           |:|||  ||||| ||:||||:|||| | |:||||||||||||||||||||||||||||||
Db     362 ANALSKLYYPPVAAVSISYPKEAIRTECLIDGELKGFGQLHPRTQGVETLGTIYSSSLFP 421

Qy     425 NRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQ 484
           |||| ||:||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||
Db     422 NRAPPGRILLLNYIGGSTNTGILSKSEGELVEAVDRDLRKMLIKPNSTDPLKLGVRVWPQ 481

Qy     485 AIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSR 544
           ||||||||| | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||
```

```
Db     482 AIPQFLVGHFDILDTAKSSLTSSGYEGLFLGGNYVAGVALGRCVEGAYETAIEVNNFMSR 541

Qy     545 YAYK 548
           ||||
Db     542 YAYK 545

RESULT   5
B70473
protoporphyrinogen oxidase - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C;Accession: B70473
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666
A;Accession: B70473
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-436 <AQF>
A;Cross-references: GB:AE000768; NID:g2984249; PIDN:AAC07778.1; PID:g2984251; GB:AE00
A;Experimental source: strain VF5
C;Genetics:
A;Gene: hemG


  Query Match             18.1%;  Score 515;  DB 2;  Length 436;
  Best Local Similarity   29.9%;  Pred. No. 2.5e-30;
  Matches  144;  Conservative  103;  Mismatches  170;  Indels   64;  Gaps   16;

Qy      73 DCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
           : |::|||||||  |  :     :::| |  || || | ||:  |||:| |  :
Db       5 EVVVIGAGISGLSTAYRLKKEGVDVVVYEKDDRIGGTIHTVKEKGYLFEVGAQTILADQE 64

Qy     133 MLTMAVDCGLKDDLVLGDPNAP-RFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGP 191
           ::    : |::   |   |::  |::  ||:| |:|    :     |:|:  ||:
Db      65 VIDFLKEAGIEP--VEASPSSKYRYIYKKGRLIPLPMSPVEFLKTPLLSLKTKLKV-LTE 121

Qy     192 IGLRPSPPGHEE--SVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKL 249
           |  |    | :|  |:  ||| : | |    :: || |||| ||| |||:| |  |:::
Db     122 IFKR----GVDEDISIADFVREHFGEEFLNYVVAPFISGVYAGDPEKLSLKHATPKLYEA 177

Qy     250 EETGGSIIGGTFKA-IKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARL--- 305
           ::  ||:|    || |||:::            |||: : || :||  | :|:: :|
Db     178 QKKYGSLI----KAFIKEKTA-----------GPKGKLI-SFGEGLGELINALAQKLEVH 221

Qy     306 --GSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVA 363
                |::  |     : :  |             ::::|:|:  |:| :| :|: :| :
Db     222 TENVVLRMR-KFEDFFRLDVRG-----------KKVETKSVVVASPAYTSSYLLKEVSFS 269

Qy     364 AADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLF 423
           |::      ||||  | :            |:|:      |   :     || :: | ||
Db     270 ASEEFDKIDYPPVVVVNVG-----------VEGKFPKVRLLSSESGKKRILGAMFMSKLF 318

Qy     424 PNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLV--VGVRV 481
           | |||:|: ||  ::||| : |::  :| ::  :|:|:|:::|     |   :  : |:
Db     319 PGRAPQGKELLTVFLGGATDREVIELSEEEIENIVERELKEIL-----QIDCIDFMHVQK 373

Qy     482 WPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGF 541
           | :||||: :|:   |: |:    |    |||| ||:: ||:   |:  : :|| :|  |
Db     374 WKRAIPQYTLGYDRFLNLAQEMEKD--YPGLFLTGNWLYGVSTADCIRASKKVAQKVLSF 431

Qy     542 L 542
           |
Db     432 L 432


RESULT   6
T04076
protoporphyrinogen oxidase (EC 1.3.3.4) IX, mitochondrial [validated] - common tobacc
C;Species: Nicotiana tabacum (common tobacco)
```

```
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C;Accession: T04076
R;Lermontova, I.; Kruse, E.; Mock, H.P.; Grimm, B.
Proc. Natl. Acad. Sci. U.S.A. 94, 8895-8900, 1997
A;Title: Cloning and characterization of a plastidal and a mitochondrial isoform of toba
A;Reference number: Z15186; MUID:97385200
A;Accession: T04076
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-504 <LER>
A;Cross-references: EMBL:Y13466; NID:g2370334; PIDN:CAA73866.1; PID:g2370335
A;Experimental source: strain SR1
C;Genetics:
A;Gene: ppxII
C;Function:
A;Description: EC 1.3.3.4 [validated; MUID:97385200]
A;Pathway: tetrapyrrole synthesis
C;Keywords: mitochondrion; oxidoreductase

  Query Match             17.8%;  Score 505;  DB 2;  Length 504;
  Best Local Similarity   28.8%;  Pred. No. 1.7e-29;
  Matches  149;  Conservative  100;  Mismatches  208;  Indels   60;  Gaps   17;

Qy        60 PSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYL 119
             ||:  |   :|:    ::|||:|||  |  :  :  |:   ||  :||| : :| :|| :
Db         3 PSAGEDKHSSAK-RVAVIGAGVSGLAAAYKLKIHGLNVTVFEAEGKAGGKLRSVSQDGLI 61

Qy       120 WEEGPNSFQPSDPMLTMAVD-CGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDL 178
             |:|| |:   |:  :|  :|  ||::          |::   |    :||   ||    :
Db        62 WDEGANTMTESEGDVTFLIDSLGLREKQQFPLSQNKRYIARNGTPVLLPSNPIDLIKSNF 121

Qy       179 MSIPGKLRAGFGPI-----GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGD 233
             :|   ||:    ||      |      | |||  | :|: | || : ||:|| :|   ||
Db       122 LSTGSKLQMLLEPILWKNKKLSQVSDSH-ESVSGFFQRHFGKEVVDYLIDPFVAGTCGGD 180

Qy       234 PSKLSMKAAFGKVWKLEETGGSIIGGTFKA-IKERSSTPKAPRDPRLPKPKGQTVGSFRK 292
             |  |||  :| ::| ||:  ||:| |  :: :  ::   : |      | : :   ||
Db       181 PDSLSMHHSFPELWNLEKRFGSVILGAIRSKLSPKNEKKQGPPKTSANKKRQRGSFSFLG 240

Qy       293 GLRMLPDAI-------SARLGSK-LKL-----------SWKLSSITKSEKGGYHLTYETP 333
             |:: | |||         || |: |:|           || : | :  ::     :::
Db       241 GMQTLTDAICKDLREDELRLNSRVLELSCSCTEDSAIDSWSIISASPHKRQSEEESFD-- 298

Qy       334 EGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY----YPPVGAVTISYPQEAIR 389
                      :::|| |     ::   ::      | ||     | |:  |  :: :| ::
Db       299 ---------AVIMTAPLCDVKSM--KIAKRGNPFLLNFIPEVDYVPLSVVITTFKRENVK 347

Qy       390 DERLVDGELKGFGQLHP---RTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEI 446
                     |:||| | |   :  |::||||::|| :||:|||    |   ::||::| |:
Db       348 ------YPLEGFGVLVPSKEQQHGLKTLGTLFSSMMFPDRAPNNVYLYTTFVGGSRNREL 401

Qy       447 LSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGH-LDTLSTAKAAMN 505
                : ::| |:|  ||:::|      :|  |    | :| |  | ||  |::  |   |
Db       402 AKASRTELKEIVTSDLKQLL--GAEGEPTYVNHLYWSKAFP--LYGHNYDSVLDAIDKME 457

Qy       506 DNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFL 542
              | | |||  ||:  |:::|: :      |  |  :|
Db       458 KN-LPGLFYAGNHRGGLSVGKALSSGCNAADLVISYL 493


RESULT    7
D47045
coproporphyrinogen III oxidase / protoporphyrinogen IX oxidase hemY - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: D47045; D69640
R;Hansson, M.; Hederstedt, L.
J. Bacteriol. 174, 8081-8093, 1992
A;Title: Cloning and characterization of the Bacillus subtilis hemEHY gene cluster, whic
A;Reference number: A47045; MUID:93094140
A;Contents: 3G18
A;Accession: D47045
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-470 <HAN>
A;Cross-references: GB:M97208; NID:g143041; PIDN:AAA22519.1; PID:g143045
A;Note: sequence extracted from NCBI backbone (NCBIN:119989, NCBIP:119993)
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter,.N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A;Reference number: A69580; MUID:98044033
A;Accession: D69640
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-470 <KUN>
A;Cross-references: GB:Z99109; GB:AL009126; NID:g2633260; PIDN:CAB12854.1; PID:e11830
A;Experimental source: strain 168
C;Genetics:
A;Gene: hemY

  Query Match             17.1%;  Score 485;  DB 2;  Length 470;
  Best Local Similarity   28.7%;  Pred. No. 4.6e-28;
  Matches  139;  Conservative   98;  Mismatches  206;  Indels   42;  Gaps   15;

Qy        75 VIVGAGISGLCIA-----QVMSANYP-NLMVTEARDRAGGNITTVERDGYLWEEGPNSF- 127
             ||:| ||:||  |     ::   | |  | : || |  || | ||::|||: | ||:||
Db         9 VIIGGGITGLAAAFYMEKEIKEKNLPLELTLVEASPRVGGKIQTVKKDGYIIERGPDSFL 68

Qy       128 --QPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLW--------KGKLRPVPSKLTDLPFFD 177
                :| | |    | || : |:: :     :||         || :  :|:|:
Db        69 ERKKSAPQLVK--DLGL-EHLLVNNATGQSYVLVNRTLHPMPKGAVMGIPTKIAPFVSTG 125

Qy       178 LMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKL 237
             | |: || ||    | |  |    ::|: :| || :| || | ||||  ||:| ||  ||
Db       126 LFSLSGKARAAMDFI-LPASKTKDDQSLGEFFRRRVGDEVVENLIEPLLSGIYAGDIDKL 184

Qy       238 SMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRML 297
             |: : | : :: |:   |:| |       : : |:        | :||   :   ||: |
Db       185 SLMSTFPQFYQTEQKHRSLILGM------KKTRPQGSGQQLTAKKQGQ-FQTLSTGLQTL 237

Qy       298 PDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNIL 357
              : |  :|  ||   :| : :||    |   : |   | |:| : |:::| |   |: :|
Db       238 VEEIEKQL--KLTKVYKGTKVTKLSHSGSCYSLELDNG-VTLDADSVIVTAPHKAAAGML 294

Qy       358 RPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTI 417
               |   |   | | :   |  | : :|: ::: |       :| | :  |
Db       295 SEL--PAISHLKNMHSTSVANVALGFPEGSVQMEH------EGTGFVISRNSDFAITACT 346

Qy       418 YSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVV 477
             :::  :|: ||:|: ||  |:| | :  |:  ::: :: :|  ||:|::      :| :
Db       347 WTNKKWPHAAPEGKTLLRAYVGKAGDESIVDLSDNDIINIVLEDLKKVM--NINGEPEMT 404

Qy       478 GVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASE 537
              |  | :::||: |||   :   : |:  :   |::: |    || :  |::      |:
Db       405 CVTRWHESMPQYHVGHKQRIKELREAL-ASAYPGVYMTGASFEGVGIPDCIDQGKAAVSD 463

Qy       538 VTGFL 542
                :|
Db       464 ALTYL 468


RESULT    8
```

```
T48595
protoporphyrinogen oxidase-like protein - Arabidopsis thaliana
N;Alternate names: protein F18O22.10
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T48595
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24493
A;Accession: T48595
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-501 <BEV>
A;Cross-references: EMBL:DB
A;Experimental source: cultivar Columbia; BAC clone F18O22
C;Genetics:
A;Map position: 5
A;Introns: 10/1; 27/2; 72/3; 94/3; 111/3; 130/3; 155/2; 167/3; 189/3; 199/3; 244/3; 298/
A;Note: F18O22.10

  Query Match             16.0%;  Score 456;  DB 2;  Length 501;
  Best Local Similarity   27.0%;  Pred. No. 7e-26;
  Matches  142;  Conservative   95;  Mismatches  194;  Indels   94;  Gaps   18;

Qy     63 AVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLWEE 122
          || |   |     :||||:|||  |  : :   |: | ||  | || : :| ::| :|:|
Db     13 AVSGKRVA-----VVGAGVSGLAAAYKLKSRGLNVTVFEADGRVGGKLRSVMQNGLIWDE 67

Qy    123 GPNSFQPSDPML-TMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSI 181
          | |:   ::| : ::  | ||::          |:::  |    :|:   :|    ::|
Db     68 GANTMTEAEPEVGSLLDDLGLREKQQFPISQKKRYIVRNGVPVMLPTNPIELVTSSVLST 127

Qy    182 PGKLRAGFGPI-----GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSK 236
            | :    |        : |    |||| :| :|: | || : ||:||  |    ||
Db    128 QSKFQILLEPFLWKKKSSKVSDASAEESVSEFFQRHFGQEVVDYLIDPFVGGTSAADPDS 187

Qy    237 LSMKAAFGKVWKLEETGGSII-----------GGTFKAIKERSSTPKAPRDPRLPKPKGQ 285
          |||| :|  :|    : ||||           ||  :  |    | |  |
Db    188 LSMKHSFPDLW---NSFGSIIVGAIRTKFAAKGGKSRDTKSSPGTKKGSRGS-------- 236

Qy    286 TVGSFRKGLRMLPDAI-------SARLGSK-LKLS---------WKLSSITKSEKGGYHL 328
             ||: |:::||| :          | || | ||         | || :: :|
Db    237 --FSFKGGMQILPDTLCKSLSHDEINLDSKVLSLSYNSGSRQENWSLSCVSHNE------ 288

Qy    329 TYETPEGVVSLQSRSIVMTVPSYVASNI--LRPLSVAAA------DALSNFYYPPVGAVT 380
                     | ::     | | |: :  :: : |         : |    | |:  :
Db    289 ----------TQRQN-----PHYDAAPLCNVKEMKVMKGGQPFQLNFLPEINYMPLSVLI 333

Qy    381 ISYPQEAIRDERLVDGELKGFGQLHP---RTQGVETLGTIYSSSLFPNRAPKGRVLLLNY 437
           :: :| ::        |:||| | |   :  | :||||::|| :||:|:|    |   :
Db    334 TTFTKEKVK------RPLEGFGVLIPSKEQKHGFKTLGTLFSSMMFPDRSPSDVHLYTTF 387

Qy    438 IGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTL 497
          |||::| |:   :  :| :||  ||:::|      :|: |    | :| | :   : |::
Db    388 IGGSRNQELAKASTDELKQVVTSDLQRLL--GVEGEPVSVNHYYWRKAFPLYDSSY-DSV 444

Qy    498 STAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFL 542
            |   | :| | | |  ||:  |:::|: :    : |  |  :|
Db    445 MEAIDKM-ENDLPGFFYAGNHRGGLSVGKSIASGCKAADLVISYL 488


RESULT   9
T07121
protoporphyrinogen oxidase (EC 1.3.3.4) IX, mitochondrial - potato (fragment)
C;Species: Solanum tuberosum (potato)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999
C;Accession: T07121
R;Johnston, D.J.
submitted to the EMBL Data Library, April 1998
A;Reference number: Z15932
A;Accession: T07121
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-404 <JOH>
A;Cross-references: EMBL:AJ225108; NID:e1289159; PIDN:CAA12401.1; PID:e1289160
A;Experimental source: cv. Bintje
C;Genetics:
A;Genome: nuclear
C;Function:
A;Pathway: tetrapyrrole synthesis
C;Keywords: mitochondrion; oxidoreductase

  Query Match             14.9%;  Score 424;  DB 2;  Length 404;
  Best Local Similarity   28.7%;  Pred. No. 1.2e-23;
  Matches  123;  Conservative   83;  Mismatches  152;  Indels   70;  Gaps   15;

Qy     60 PSSAVDGGPAAELDC----VIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVER 115
          ||:  |     : :|     ::|||:|||  |  :  :  |: | ||  |||| : :: :
Db      3 PSAGED-----KQNCPKRVAVIGAGVSGLAAAYKLKIHGLNVTVFEAEGRAGGKLRSLSQ 57

Qy    116 DGYLWEEGPNSFQPSDPMLTMAVD-CGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLP 174
          || :|:|| |:   |:  :|  :|  ||::          |::   |    :||   ||
Db     58 DGLIWDEGANTMTESEGDVTFLLDSLGLREKQQFPLSQNKRYIARNGTPTLIPSNPIDLI 117

Qy    175 FFDLMSIPGKLRAGFGPIGLRPSP----PGHEESVEQFVRRNLGGEVFERLIEPFCSGVY 230
            : :|   ||:  | |:  : :         |||  | :|: | || : ||:|| :|
Db    118 KSNFLSTGSKLQMLFEPLLWKNNKLTKVSDEHESVSGFFQRHFGKEVVDYLIDPFVAGTC 177

Qy    231 VGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKA----IKERSSTPKAPRDPRLPKPKGQT 286
           |||  |||  :| ::| ||:  ||:| |  ::    |||:   |     |:    | :
Db    178 GGDPDSLSMHLSFPELWNLEKRFGSVIVGAIRSKLSPIKEKKQGP-----PKTSVNKKRQ 232

Qy    287 VGSFR--KGLRMLPDAI-------SARLGSK-LKL-----------SWKLSSITKSEKGG 325
           |||    |:: | |||         || |: |:|           || : | :  ::
Db    233 RGSFSFLGGMQTLTDAICKDLKEDELRLNSRVLELSCSCSGDSAIDSWSIFSASPHKRQA 292

Qy    326 YHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY----YPPVGAVTI 381
             :::           :::|| |     ::   ::      | ||     | |:  |
Db    293 EEESFD-----------AVIMTAPLCDVKSM--KIAKRGNPFLLNFIPEVDYVPLSVVIT 339

Qy    382 SYPQEAIRDERLVDGELKGFGQLHP---RTQGVETLGTIYSSSLFPNRAPKGRVLLLNYI 438
          :: :|:::        |:||| | |   :  |::||||::|| :||:|||    |   ::
Db    340 TFKKESVKH------PLEGFGVLVPSQEQKHGLKTLGTLFSSMMFPDRAPNNVYLYTTFV 393

Qy    439 GGAKNPEI 446
          ||::| |:
Db    394 GGSRNREL 401


RESULT  10
D83800
protoporphyrinogen IX and coproporphyrinogen III oxidase hemY [imported] - Bacillus h
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
C;Accession: D83800
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A;Reference number: A83650; MUID:20263314
A;Accession: D83800
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <STO>
A;Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BAB04923.1; GSPDB:G
A;Experimental source: strain C-125
C;Genetics:
A;Gene: hemY

  Query Match             14.3%;  Score 407.5;  DB 2;  Length 467;
  Best Local Similarity   28.2%;  Pred. No. 2.4e-22;
  Matches  135;  Conservative   90;  Mismatches  210;  Indels   43;  Gaps   19;

Qy     76 IVGAGISGLCIAQVMSANYPNLMV----TEARDRAGGNITTVERDGYLWEEGPNSF-QPS 130
```

```
            ::| |::||  |  :      : :     ||  : || | |   :|:: | ||:|:
Db       10 VIGGGMTGLAAAWALQKARKQMPIEYELIEASAQLGGKIQTDVTEGFVIERGPDSYLARK 69

Qy      131 DPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKL-----TDL-PFF--DLMSIP 182
            : |  :| : ||: :||  |     ::|   :| |:|        |:: ||    | |
Db       70 ESMTRLAREVGLEGELVRND-TGQAYILKDDQLYPIPGGAIMGIPTEIGPFIRSRLFSSI 128

Qy      183 GKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAA 242
            |||||    :    :    : |:   | |: || ||  ::||||  ||:| ||   ||::|
Db      129 GKLRAAQDLVIPSFTHENEDVSLGYFFRKRLGDEVVDQLIEPLLSGIYAGDLDTLSLQAT 188

Qy      243 FGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAIS 302
            | :  :||   ||:| |  || | : |        :: | :|  : :||:||:   : |
Db      189 FPQFQQLEAKYGSLIRG-IKASKGKQS------QSQVGKKQGMFL-TFRRGLQSFIETIE 240

Qy      303 ARLGSKLKLSWKLSSITKSEKGG--YHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360
              |    |   | : :   |: |  | ||:   :| |      :::| | ||   :| |
Db      241 QHLDE--KAIRKKTEVRFCEREGKRYRLTFN--DGTVQTYDH-VILTTPPYVTGRLLEPY 295

Qy      361 SVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420
                   :   |   : | |::    | : | |     :| | : ||    |     ::
Db      296 -------VDVDYMREMKATTVATVAMAFKSEDL-HIPYEGTGFVVPRKSRYEITACTWTH 347

Qy      421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRK-MLIKPKAQDPLVVGV 479
              : :  |:|  |:  |:|   :  |: :|:  :|| | :|| : | :|   ::||   |
Db      348 KKWRHTTPEGHALVRCYVGKPGSESIVHETDEAIVEQVQKDLDQIMTVK---REPLFYKV 404

Qy      480 RVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCV-EGAYEVAS 536
              | :|:||: |||   |   :  :    : |: | |    |: |  |: :|   ||:
Db      405 TRWKKAMPQYEVGHRYKLEKVRQELK-KAMPGIILAGAGFDGIGLPDCIDQGEAAVAT 461

RESULT  11
S65684
protoporphyrinogen oxidase - mouse
C;Species: Mus musculus (house mouse)
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: S65684
R;Taketani, S.; Yoshinaga, T.; Furukawa, T.; Kohno, H.; Tokunaga, R.; Nishimura, K.; Inc
Eur. J. Biochem. 230, 760-765, 1995
A;Title: Induction of terminal enzymes for heme biosynthesis during differentiation of m
A;Reference number: S65629; MUID:95331315
A;Accession: S65684
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-477 <TAK>
A;Cross-references: EMBL:D45185; NID:g1060929; PIDN:BAA08126.1; PID:g1786169
C;Genetics:
A;Genome: nuclear
C;Superfamily: phytoene dehydrogenase
C;Keywords: mitochondrion


  Query Match             13.0%;  Score 368.5;  DB 2;  Length 477;
  Best Local Similarity   25.0%;  Pred. No. 1.9e-19;
  Matches  134;  Conservative   75;  Mismatches  188;  Indels  139;  Gaps   19;

Qy       75 VIVGAGISGLCIAQ--VMSANYPNLMVTEARDRAGGNITTVE-RDGYLWEEGPNSFQPSD 131
            :::| |||||  :   :   : | ::: |   | || | ::   || ::| ||   :|:
Db        6 IVLGGGISGLAASYHLIRGPSPPKVILVEGSKRLGGWIRSIRGSDGAIFELGPRGIRPTG 65

Qy      132 PM----LTMAVDCGLKDDL--VLGDPNAP--RFVLWKGKLRPVPSKLTDLPFFDLMSIPG 183
            |:    | :  : ||: ::  | ||  |   ||:   | | | || |  |
Db       66 PLGARTLLLVSELGLESEVLPVRGDHPAAQNRFLYVGGTLHPSPSGLRGL---------- 115

Qy      184 KLRAGFGPIGLRPSPP-----------------GHE--ESVEQFVRRNLGGEVFERLIEP 224
                      ||||||                 | |  |:|  | :| || ||    ::
Db      116 ----------LRPSPPFSKPLFWAGLRELLKPRGKEPDETVHSFAQRRLGPEVASLAMDS 165

Qy      225 FCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKG 284
             | ||: |:  :||::: |  ::: |:|  ||: |      :             |:|
Db      166 LCRGVFAGNSRELSIRSCFPSLFQAEQTHRSILLGLLLGAGQS------------PQPDS 213
```

```
Qy      285 QTV----------GSFRKGLRMLPDAISARLGSK------------LKLS----WKLSSI 318
              :           | | || :|| |:   | ||            | |     ||:|
Db      214 SLIRQARAERWSQWSLRGGLEVLPQALHNHLASKGVTVLSGQPVCGLSLQPEGRWKVSL- 272

Qy      319 TKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGA 378
                            |  ||::  |:  :|:   | :|   :   |  ||      |
Db      273 ----------------GDSSLEADHIISAIPASELSKLLPAEAAPLARILSTIKAVSVAV 316

Qy      379 VTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNR---APKGRVLLL 435
            | : |    :         ::||| | | ::    || :| |  || :    |  || ::
Db      317 VNLQYRGACL--------PVQGFGHLVPSSEDPTVLGIVYDSVAFPEQDGNPPSLRVTVM 368

Qy      436 NYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQD-----------PLVVGVRVWPQ 484
              :||            | ::     |   | : :||:           |    | :
Db      369 --LGG---------YWLQKLKAAGHQLSPELFQQQAQEAAATQLGLKEPPSHCLVHLHKN 417

Qy      485 AIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540
             |||: :||   | :|   :    |     | :|  |||:  |:|   : |  | |
Db      418 CIPQYTIGHCQKLDSAMQFLTAQRLPLTLAGASY-EGVAVNDCIESGRQAAVAVLG 472

RESULT  12
A75435
protoporphyrinogen oxidase - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C;Accession: A75435
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896
A;Accession: A75435
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-462 <WHI>
A;Cross-references: GB:AE001962; GB:AE000513; NID:g6458855; PIDN:AAF10699.1; PID:g645
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1130
A;Map position: 1


  Query Match             12.8%;  Score 363.5;  DB 2;  Length 462;
  Best Local Similarity   27.4%;  Pred. No. 4.2e-19;
  Matches  141;  Conservative   85;  Mismatches  186;  Indels  103;  Gaps   23;

Qy       65 DGGP-----AAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVER--DG 117
            :|||     |  |  ::||||::||  |  :       :: || |  || : : ||  |
Db       10 EGGPMTNPNAHTLPIIVVGAGMTGLAAAWELQQRGVPYVLLEAGDSFGGKVQS-ERTEDE 68

Qy      118 YLWEEGPNSFQPSDP-MLTMAVDCGLKDDLVLGDPNAPR---FVLWKGKLRPVPSKLTDL 173
            :| |:  ::|    |  | :| : ||:|:|:   | |     : |  |:|   |    :|
Db       69 FLVEKAADAFILGKPWALQLAQEVGLQDELI--HPRAETKKLYFLRGGELLDFPQ---NL 123

Qy      174 PFFDLMSIPGKLRAG-FGPIGLRPS------PP----GHEESVEQFVRRNLGGEVFERLI 222
              |  :     |::|   | |||        ||      :||:  |: |  | |     |
Db      124 KMFVPLDDESFLQSGVLSPEGLRRMLDEQNVPPKENTDEDESLADFIIRRFGAEAM-NFI 182

Qy      223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
             |  :|:|| :| :||||||| :   ||:  ||:| |:        :||:|
Db      183 VPLAAGIYVANPFELSMKAAFPQFLALEQKYGSVIKGS-------RATPRA--------- 226

Qy      283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVV----- 337
            :|   |||: |:  | :|::|:|   ::|:  : ::              |:||
Db      227 QGPIFGSFKDGMSTLANAVAAKLTGDVRLNTPVLAV-------------HPDGVTLTGGE 273

Qy      338 SLQSRSIVMTVPSYVASNILRPLSVAAADALSNFYYPPVG--------AVTISYPQ-EAI 388
             : :  :|:| |::||:  |   |  || ||       ||        || ::| : :
Db      274 KIAASGVVLTAPAWVAAPTLGQ-SFPAASAL-------VGELQTNGSVAVILAYREGQFP 325

Qy      389 RDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILS 448
```

```
           ||  |        |      :|:|       |:    |||:| |||  :       : |:
Db     326 RDMHL-------HGLQVAADEGIEMKAITVHSAKLHGRAPEGHVLLRVFFKDLDPQKALA 378

Qy     449 KTESQLVEVVDRDLRKMLIKPKAQ-DPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDN 507
           :   ::  :             || |||      |    | : ||||| ::  :||:  |
Db     379 EAVHEVARLFG-----------AQGDPLWHAYADWRGKNPAYQVGHLDHIARIRAALPPN 427

Qy     508 GLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFL 542
               : : |:  :|| :  ||      | :|   |
Db     428 ----IQVAGSSYTGVGIPDCVNAGRTAARDVVAAL 458


RESULT  13
A56449
protoporphyrinogen oxidase (EC 1.3.3.4) - human
C;Species: Homo sapiens (man)
C;Date: 27-Oct-1995 #sequence_revision 27-Oct-1995 #text_change 20-Jun-2000
C;Accession: JC4971; A56449; I38995
R;Puy, H.; Robreau, A.M.; Rosipal, R.; Nordmann, Y.; Deybach, J.C.
Biochem. Biophys. Res. Commun. 226, 226-230, 1996
A;Title: Protoporphyrinogen oxidase: Complete genomic sequence and polymorphisms in the
A;Reference number: JC4971; MUID:96400242
A;Accession: JC4971
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-477 <PUY>
A;Cross-references: EMBL:X99450
R;Nishimura, K.; Taketani, S.; Inokuchi, H.
J. Biol. Chem. 270, 8076-8080, 1995
A;Title: Cloning of a human cDNA for protoporphyrinogen oxidase by complementation in Vi
A;Reference number: A56449; MUID:95229621
A;Accession: A56449
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-477 <NIS>
A;Cross-references: GB:D38537; NID:g854663; PIDN:BAA07538.1; PID:g854664
R;Dailey, T.A.; Meissner, P.; Dailey, H.A.
J. Biol. Chem. 269, 813-815, 1994
A;Title: Expression of a cloned protoporphyrinogen oxidase.
A;Reference number: A49858; MUID:94117488
A;Accession: I38995
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-477 <RES>
A;Cross-references: EMBL:U26446; NID:g837327; PIDN:AAA67690.1; PID:g837328
C;Comment: This enzyme acts as the penultimate step in the heme biosynthetic pathway, an
.
C;Genetics:
A;Gene: GDB:PPOX; PPO
A;Cross-references: GDB:118852; OMIM:600923
A;Map position: 1q23-1q23
A;Introns: 29/3; 74/3; 113/2; 157/3; 206/1; 269/3; 290/1; 329/3; 366/3; 416/3; 431/1
C;Superfamily: phytoene dehydrogenase
C;Keywords: oxidoreductase


  Query Match             12.8%;  Score 363.5;  DB 2;  Length 477;
  Best Local Similarity   25.5%;  Pred. No. 4.4e-19;
  Matches  138;  Conservative   83;  Mismatches  170;  Indels  151;  Gaps   23;

Qy      75 VIVGAGISGLCIAQVMSANY--------PNLMVTEARDRAGGNITTVE-RDGYLWEEGPN 125
           |::| |||||      :|:|        | ::: |: :| || | :|   :| ::| ||
Db       6 VVLGGGISGL------AASYHLSRAPCPPKVVLVESSERLGGWIRSVRGPNGAIFELGPR 59

Qy     126 SFQPSDPM----LTMAVDCGLKDDL--VLGDPNAP--RFVLWKGKLRPVPSKLTDLPFFD 177
             :|:  :    | :  : ||  ::  | ||  |   ||:   | |  :|: |  |
Db      60 GIRPAGALGARTLLLVSELGLDSEVLPVRGDHPAAQNRFLYVGGALHALPTGLRGL---- 115

Qy     178 LMSIPGKLRAGFGPIGLRPSPP-----------------GHE--ESVEQFVRRNLGGEVF 218
                           ||||||                 | |  |:|  | :| || ||
Db     116 ----------------LRPSPPFSKPLFWAGLRELTKPRGKEPDETVHSFAQRRLGPEVA 159

Qy     219 ERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPR 278
```

```
              ::  | ||: |:  :||::: |  ::: |:|  ||: |                  |
Db     160 SLAMDSLCRGVFAGNSRELSIRSCFPSLFQAEQTHRSILLGLLLGA-----------GR 207

Qy     279 LPKPKGQTV----------GSFRKGLRMLPDAISARLGSK------------LKLS---- 312
            |:|    :           | | || ||| |:   | |:            | |
Db     208 TPQPDSALIRQALAERWSQWSLRGGLEMLPQALETHLTSRGVSVLRGQPVCGLSLQAEGR 267

Qy     313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
           ||: |:  |                ||::  ::  :|: | | :|   :   | |||
Db     268 WKV-SLRDS----------------SLEADHVISAIPASVLSELLPAEAAPLARALSAIT 310

Qy     373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNR--APKG 430
              |  | : |    :         ::||| | | ::     || :| |  || :  :| |
Db     311 AVSVAVVNLQYQGAHL--------PVQGFGHLVPSSEDPGVLGIVYDSVAFPEQDGSPPG 362

Qy     431 -RVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLV--VGVRVWP---- 483
            || ::  :||         :  | :|     | : | : :||:     :|::  |
Db     363 LRVTVM--LGG---------SWLQTLEASGCVLSQELFQQRAQEAAATQLGLKEMPSHCL 411

Qy     484 -----QAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEV 538
                  |||: :||   | :|:  :  : |     | :|  |||:  |:|   : |  |
Db     412 VHLHKNCIPQYTLGHWQKLESARQFLTAHRLPLTLAGASY-EGVAVNDCIESGRQAAVSV 470

Qy     539 TG 540
            |
Db     471 LG 472


RESULT  14
S68367
protoporphyrinogen oxidase (EC 1.3.3.4) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 19-May-2000
C;Accession: S68367; S65630
R;Dailey, T.A.; Dailey, H.A.; Meissner, P.; Prasad, A.R.K.
Arch. Biochem. Biophys. 324, 379-384, 1995
A;Title: Cloning, sequence, and expression of mouse protoporphyrinogen oxidase.
A;Reference number: S68367; MUID:96132666
A;Accession: S68367
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-477 <DAI>
A;Cross-references: EMBL:U25114; NID:g793939; PIDN:AAA96003.1; PID:g793940
R;Taketani, S.; Yoshinaga, T.; Furukawa, T.; Kohno, H.; Tokunaga, R.; Nishimura, K.;
Eur. J. Biochem. 230, 760-765, 1995
A;Title: Induction of terminal enzymes for heme biosynthesis during differentiation o
A;Reference number: S65629; MUID:95331315
A;Accession: S65630
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-14;164-178 <TAK>
C;Genetics:
A;Genome: nuclear
C;Superfamily: phytoene dehydrogenase
C;Keywords: mitochondrion; oxidoreductase


  Query Match             12.7%;  Score 362.5;  DB 2;  Length 477;
  Best Local Similarity   24.8%;  Pred. No. 5.2e-19;
  Matches  133;  Conservative   76;  Mismatches  188;  Indels  139;  Gaps   19;

Qy      75 VIVGAGISGLCIAQ--VMSANYPNLMVTEARDRAGGNITTVE-RDGYLWEEGPNSFQPSD 131
           :::| |||||  :   :   : | ::: |   | || |  ::   || ::| ||   :|:
Db       6 IVLGGGISGLAASYHLIRGPSPPKVILVEGSKRLGGWIRSIRGSDGAIFELGPRGIRPAG 65

Qy     132 PM----LTMAVDCGLKDDL--VLGDPNAP--RFVLWKGKLRPVPSKLTDLPFFDLMSIPG 183
            :    | :  : ||: ::  | ||  |   ||:   | | |:|| |  |
Db      66 ALGARTLLLVSELGLESEVLPVRGDHPAAQNRFLYVGGTLHPLPSGLRGL---------- 115

Qy     184 KLRAGFGPIGLRPSPP-----------------GHE--ESVEQFVRRNLGGEVFERLIEP 224
                     ||||||                 | |  |:|  | :| || ||    ::
Db     116 ----------LRPSPPFSKPLFWAGLRELLKPRGKEPDETVHSFAQRRLGPEVASLAMDS 165
```

```
Qy     225 FCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKG 284
            | ||: |:  :||::: |  ::: |:|  ||: |      :             |:|
Db     166 LCRGVFAGNSRELSIRSCFPSLFQAEQTHRSILLGLLLGAGQS------------PQPDS 213

Qy     285 QTV----------GSFRKGLRMLPDAISARLGSK------------LKLS----WKLSSI 318
             :           | | || :|| |:   | ||            | |     ||:|
Db     214 SLIRQARAERWSQWSLRGGLEVLPQALHNHLASKGVTVLSGQPVCGLSLQPEGRWKVSL- 272

Qy     319 TKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGA 378
                           |  ||::  |:  :|:   | :|   :   |  ||      |
Db     273 ----------------GDSSLEADHIISAIPASELSKLLPAEAAPLARILSTIKAVSVAV 316

Qy     379 VTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNR---APKGRVLLL 435
           | : |    :         ::||| | | ::    || :| |  || :    |  || ::
Db     317 VNLQYRGACL--------PVQGFGHLVPSSEDPTVLGIVYDSVAFPEQDGNPPSLRVTVM 368

Qy     436 NYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQD-----------PLVVGVRVWPQ 484
             :||            | ::     |   | : :||:           |    | :
Db     369 --LGG---------YWLQKLKAAGHQLSPELFQQQAQEAAATQLGLKEPPSHCLVHLHKN 417

Qy     485 AIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540
            |||: :||   |  :|   :    |       :|  |||:  |:|   : |  | |
Db     418 CIPQYTIGHWQKLDSAMQFLTAQRLPLTLAGASY-EGVAVNDCIESGRQAAVAVLG 472

RESULT  15
T34679
probable protoporphyrinogen oxidase - Streptomyces coelicolor (fragment)
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 03-Nov-2000
C;Accession: T34679
R;Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z21553
A;Accession: T34679
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-468 <HAR>
A;Cross-references: EMBL:AL023517; PIDN:CAA18975.1; GSPDB:GN00070; SCOEDB:SC1B5.01
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC1B5.01
C;Superfamily: phytoene dehydrogenase


  Query Match             11.1%;  Score 317;  DB 2;  Length 468;
  Best Local Similarity   26.5%;  Pred. No. 1.2e-15;
  Matches  129;  Conservative   77;  Mismatches  217;  Indels   64;  Gaps   17;

Qy      81 ISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP-MLTMAVD 139
           |:||  |  :      : | ||  | || :   |  |   : |  |     |  : :|
Db       1 IAGLAAAHRLLEAGARVTVLEASGRVGGKLLPGEIAGVRVDLGAESMLARRPEAVGLARA 60

Qy     140 CGLKDDLVLGDPNAPRFVLW-KGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPIG----- 193
            || |   |  |:     || :| |||:|          :| :||   |  | :
Db      61 AGLADR--LQPPSTATASLWTRGALRPMPKG-------HVMGVPGTAAALSGVLSEEGLA 111

Qy     194 -------LRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKV 246
                  |  :  | : :| ::|   || || :||:||   ||| ||  ::|:::|  ::
Db     112 RIERDAELPRTEVGDDVAVGEYVAARLGREVVDRLVEPLLGGVYAGDAYRISLRSAVPQL 171

Qy     247 WKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLP----DAIS 302
           ::   |  |:     :|:: |::|           | |        |:  ||    |::
Db     172 FEAARTHTSLTEAV-RALQGRTAT---------SPPSGPVFMGIEGGIGTLPPAVADSVR 221

Qy     303 ARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSV 362
           || | ::     :: : ::   |: :      |   | : ::|: ||:  |: :||  :
Db     222 AR-GGEIVTQAPVTELRRTASDGWRIV----AGDRVLHAGAVVVAVPAGPAAELLRAEAP 276

Qy     363 AAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSL 422
           |||  ||   |  :  ||::|     | |     |  ||  | |   |     ||  :|
Db     277 AAAAELSAVEYASMALVTLAYR----RSEAAALPEGSGF--LVPPVDG----HTI-KAST 325
```

```
Qy     423 FPNR-------APKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPL 475
           | :|            |:|   :|   : ||| : :: || |   || :      |  |:
Db     326 FASRKWGWIADEDPDLVVLRTSVGRYGDTEILGRDDAGLVAVSRHDLAEATGLTAA--PV 383

Qy     476 VVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVA 535
              |  |   :||: |||      |:   :   | || : |    || :   :  ||  |
Db     384 ATRVTRWQDGLPQYPVGH--HARVARVREHVAKLPGLAVCGAAYDGVGIPASIASAYAAA 441

Qy     536 SEVTGFL 542
            :: | |
Db     442 DQIRGDL 448


Search completed: July  3, 2001, 10:34:11
Job time: 374 sec
```

```
                  GenCore version 4.5
         Copyright (c) 1993 - 2000  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         July  3, 2001, 10:00:01 ; Search time 49.77 Seconds
                                          (without alignments)
                                          667.509 Million cell updates/sec

Title:          US-09-508-418-2
Perfect score:  2844
Sequence:       1 MTTTPIANHPNIFTHQSSSS.........EGAYEVASEVTGFLSRYAYK 548

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:      412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
                1:  /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
                2:  /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
                3:  /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
                4:  /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
                5:  /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
                6:  /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
                7:  /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
                8:  /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
                9:  /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
                10:  /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
                11:  /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
                12:  /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
                13:  /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
                14:  /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
                15:  /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
                16:  /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
                17:  /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
                18:  /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
                19:  /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
                20:  /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                21:  /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
                22:  /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                               SUMMARIES
                     %
Result              Query
  No.    Score   Match Length DB  ID                        Description
-------------------------------------------------------------------------
   1     2836    99.7    548  20  AAY05200                  N. tabacum protopo
   2     2836    99.7    548  20  AAY05201                  N. tabacum protopo
   3     2113    74.3    539  18  AAW41608                  Soybean protox-1.
   4     2113    74.3    539  18  AAW41609                  Cotton protox-1.
   5     2113    74.3    539  18  AAW25740                  Cotton protoporphy
   6     2113    74.3    539  22  AAB72907                  Cotton protoporphy
   7     2108    74.1    539  22  AAB72921                  Cotton protoporphy
   8     2104    74.0    539  22  AAB72922                  Cotton protoporphy
   9     2078    73.1    537  22  AAB72923                  Arabidopsis protop
  10     2075    73.0    537  17  AAR90295                  Protein having pro
  11     2075    73.0    537  18  AAW41603                  Arabidopsis thalia
```

```
  12     2075    73.0    537  18  AAW25746                  Arabidopsis protop
  13     2075    73.0    537  19  AAW51347                  Arabidopsis proto-
  14     2075    73.0    537  21  AAG42385                  Arabidopsis thalia
  15     2075    73.0    537  22  AAB72901                  Arabidopsis protop
  16     2074    72.9    537  22  AAB72924                  Arabidopsis protop
  17     2066    72.6    537  22  AAB72925                  Arabidopsis protop
  18     2055    72.3    543  18  AAW25739                  Soybean protoporph
  19     2055    72.3    543  20  AAW92498                  Glycine max protop
  20     2055    72.3    543  21  AAY54100                  The chloroplast-ty
  21     2055    72.3    543  22  AAB72906                  Soybean protoporph
  22   2006.5    70.6    536  18  AAW41611                  Rape protox-1.  Br
  23   2006.5    70.6    536  18  AAW25742                  Rape protoporphyri
  24   2006.5    70.6    536  22  AAB72909                  Oilseed rape proto
  25     1987    69.9    528  18  AAW41607                  Wheat protox-1.  T
  26     1987    69.9    528  18  AAW25738                  Wheat protoporphyr
  27     1987    69.9    528  22  AAB72905                  Wheat protoporphyr
  28     1978    69.5    481  18  AAW41605                  Maize protox-1.  Z
  29     1978    69.5    481  18  AAW25745                  Maize protoporphyr
  30     1978    69.5    481  19  AAW51255                  Maize proto-porphy
  31     1978    69.5    481  22  AAB72903                  Maize protoporphyr
  32     1978    69.5    483  17  AAR90297                  Maize protox-1 par
  33     1972    69.3    481  22  AAB72926                  Maize protoporphyr
  34     1968    69.2    481  22  AAB72920                  Maize protoporphyr
  35   1933.5    68.0    560  18  AAW41610                  Sugar beet protox-
  36   1933.5    68.0    560  18  AAW25741                  Sugarbeet protopor
  37   1933.5    68.0    560  22  AAB72908                  Sugar beet protopo
  38     1823    64.1    440  18  AAW41613                  Sorghum protox-1.
  39     1823    64.1    440  18  AAW25744                  Sorghum protoporph
  40     1823    64.1    440  22  AAB72911                  Sorghum protoporph
  41     1786    62.8    416  21  AAG42386                  Arabidopsis thalia
  42     1772    62.3    413  21  AAG42387                  Arabidopsis thalia
  43   1530.5    53.8    563  20  AAY49538                  Chlamydomonas rein
  44   1530.5    53.8    563  21  AAY54101                  An algal protoporp
  45     1261    44.3    312  18  AAW41612                  Rice protox-1.  Or

                               ALIGNMENTS

RESULT   1
AAY05200
ID   AAY05200 standard; Protein; 548 AA.
XX
AC   AAY05200;
XX
DT   16-JUN-1999  (first entry)
XX
DE   N. tabacum protoporphyrinogen oxidase.
XX
KW   Protoporphyrinogen oxidase; light-requiring herbicide tolerance; enzyme.
XX
OS   Nicotiana tabacum.
XX
PN   WO9913087-A1.
XX
PD   18-MAR-1999.
XX
PF   10-SEP-1998;   98WO-JP04064.
XX
PR   11-SEP-1997;   97JP-0265084.
XX
PA   (NIPY ) NIPPON NOYAKU CO LTD.
XX
PI   Hirooka T,  Horikoshi M,  Mametsuka K;
XX
DR   WPI; 1999-215067/18.
DR   N-PSDB; AAX28278.
XX
PT   Protoporphyrinogen oxidase tolerant to light-requiring herbicides,
PT   useful for constructing plants with such properties
XX
PS   Claim 1; Page 34-39; 56pp; Japanese.
XX
```

```
CC   This sequence is the Nicotiana tabacum protoporphyrinogen oxidase
CC   of the invention, which is tolerant to light-requiring herbicides. The
CC   enzyme and its derivatives can be applied to construct plants with high
CC   tolerance to light-requiring herbicides, useful in agriculture.
XX
SQ   Sequence   548 AA;

  Query Match             99.7%;  Score 2836;  DB 20;  Length 548;
  Best Local Similarity   99.8%;  Pred. No. 8.7e-264;
  Matches  547;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 mtttpianhpnifthqssssplaflnrtsfipfssiskrnsvncngwrtrcsvakdytvp 60

Qy       61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 ssavdggpaaeldcvivgagisglciaqvmsanypnlmvteardraggnittverdgylw 120

Qy      121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 eegpnsfqpsdpmltmavdcglkddlvlgdpnaprfvlwkgklrpvpskltdlpffdlms 180

Qy      181 IPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMK 240
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 ipgklragfgaiglrpsppgheesveqfvrrnlggevferliepfcsgvyvgdpsklsmk 240

Qy      241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 aafgkvwkleetggsiiggtfkaikersstpkaprdprlpkpkgqtvgsfrkglrmlpda 300

Qy      301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 isarlgsklklswklssitksekggyhltyetpegvvslqsrsivmtvpsyvasnilrpl 360

Qy      361 SVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 svaaadalsnfyyppvgavtisypqeairderlvdgelkgfgqlhprtqgvetlgtiyss 420

Qy      421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 slfpnrapkgrvlllnyiggaknpeilsktesqlvevvdrdlrkmlikpkaqdplvvgvr 480

Qy      481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 vwpqaipqflvghldtlstakaamndngleglflggnyvsgvalgrcvegayevasevtg 540

Qy      541 FLSRYAYK 548
              ||||||||
Db      541 flsryayk 548

RESULT    2
AAY05201
ID   AAY05201 standard; Protein; 548 AA.
XX
AC   AAY05201;
XX
DT   16-JUN-1999  (first entry)
XX
DE   N. tabacum protoporphyrinogen oxidase mutant.
XX
KW   Protoporphyrinogen oxidase; light-requiring herbicide tolerance; enzyme.
XX
OS   Nicotiana tabacum.
XX
PN   WO9913087-A1.
XX
PD   18-MAR-1999.
XX
PF   10-SEP-1998;   98WO-JP04064.
XX
PR   11-SEP-1997;   97JP-0265084.
XX
PA   (NIPY ) NIPPON NOYAKU CO LTD.
XX
PI   Hirooka T,  Horikoshi M,  Mametsuka K;
XX
DR   WPI; 1999-215067/18.
DR   N-PSDB; AAX28279.
XX
PT   Protoporphyrinogen oxidase tolerant to light-requiring herbicides,
PT   useful for constructing plants with such properties
XX
PS   Example 5; Page 44-49; 56pp; Japanese.
XX
CC   This sequence is the Nicotiana tabacum protoporphyrinogen oxidase
CC   of the invention, which is tolerant to light-requiring herbicides. The
CC   enzyme and its derivatives can be applied to construct plants with high
CC   tolerance to light-requiring herbicides, useful in agriculture.
XX
SQ   Sequence   548 AA;

  Query Match             99.7%;  Score 2836;  DB 20;  Length 548;
  Best Local Similarity   99.8%;  Pred. No. 8.7e-264;
  Matches  547;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNCNGWRTRCSVAKDYTVP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 mtttpianhpnifthqssssplaflnrtsfipfssiskrnsvncngwrtrcsvakdytvp 60

Qy       61 SSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEARDRAGGNITTVERDGYLW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 ssavdggpaaeldcvivgagisglciaqvmsanypnlmvteardraggnittverdgylw 120

Qy      121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 eegpnsfqpsdpmltmavdcglkddlvlgdpnaprfvlwkgklrpvpskltdlpffdlms 180

Qy      181 IPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMK 240
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 ipgklragfgaiglrpsppgheesveqfvrrnlggevferliepfcsgvyvgdpsklsmk 240

Qy      241 AAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 aafgkvwkleetggsiiggtfkaikersstpkaprdprlpkpkgqtvgsfrkglrmlpda 300

Qy      301 ISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 isarlgsklklswklssitksekggyhltyetpegvvslqsrsivmtvpsyvasnilrpl 360

Qy      361 SVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 svaaadalsnfyyppvgavtisypqeairderlvdgelkgfgqlhprtqgvetlgtiyss 420

Qy      421 SLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVR 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 slfpnrapkgrvlllnyiggaknpeilsktesqlvevvdrdlrkmlikpkaqdplvvgvr 480

Qy      481 VWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 vwpqaipqflvghldtlstakaamndngleglflggnyvsgvalgrcvegayevasevtg 540

Qy      541 FLSRYAYK 548
              ||||||||
Db      541 flsryayk 548

RESULT    3
AAW41608
ID   AAW41608 standard; Protein; 539 AA.
XX
AC   AAW41608;
XX
```

```
DT   20-APR-1998  (first entry)
XX
DE   Soybean protox-1.
XX
KW   Protoporphyrinogen oxidase-1; protox-1; promoter; soybean;
KW   herbicide resistance; breeding programme; probe; gene isolation;
KW   genomic mapping.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             55..1686
FT                   /*tag=  a
FT                   /product= protox-1
XX
PN   WO9732028-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US03343.
XX
PR   21-JUN-1996;   96US-0020003.
PR   28-FEB-1996;   96US-0012705.
PR   28-FEB-1996;   96US-0013612.
XX
PA   (NOVS ) NOVARTIS AG.
XX
PI   Johnson MA,  Volrath SL,  Ward ER;
XX
DR   WPI; 1997-489209/45.
DR   N-PSDB; AAV04309.
XX
PT   DNA containing a plant proto-porphyrinogen oxidase gene promoter -
PT   optionally linked to a heterologous gene, especially to express
PT   herbicide-resistant enzymes, and plants containing such constructs
XX
PS   Claim 33; Pages 68-71; 114pp; English.
XX
CC   The present sequence is soybean protoporphyrinogen oxidase-1
CC   (protox-1).
CC   The protox-1 promoter can be used to express herbicide resistant
CC   enzymes, specifically protox, i.e. a plant tissue, plant or progeny
CC   containing a chimeric gene of the promoter and a heterologous
CC   coding sequence. The plant can also be used in breeding programmes.
CC   Also hybridising fragments of the protox coding sequence can be
CC   used as probes, e.g. to isolate related genes or for genomic
CC   mapping.
XX
SQ   Sequence   539 AA;

  Query Match             74.3%;  Score 2113;  DB 18;  Length 539;
  Best Local Similarity   79.2%;  Pred. No. 2.8e-194;
  Matches  401;  Conservative   46;  Mismatches   55;  Indels    4;  Gaps    1;

Qy       47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
            :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db       34 fklrcslaegptissskidggessiadcvivgggisglciaqalatkhrdvasnvivtea 93

Qy      103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
            ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db       94 rdrvggnittverdgylweegpnsfqpsdpiltmavdsglkddlvlgdpnaprfvlwegk 153

Qy      163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
            ||||||| ||||||||||| |||||||| ||:|| |||:|||||:||||||| ||||| |
Db      154 lrpvpskptdlpffdlmsiagklragfgaigirppppgyeesveefvrrnlgaevferfi 213

Qy      223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
            |||||||| |||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db      214 epfcsgvyagdpsklsmkaafgrvwkleeiggsiiggtfktiqernktpkpprdprlpkp 273

Qy      283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
            |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
```

```
Db      274 kgqtvgsfrkgltmlpeaianslgsnvklswklssitklgnggynltfetpegmvslqsr 333

Qy      343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
            |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db      334 svvmtipshvasnllhplsaaaadalsqfyyppvasvtvsypkeairkeclidgelkgfg 393

Qy      403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
            |||||:||:||||||||||||||||| |||||||||||| |  |||||| :||| |||||
Db      394 qlhprsqgietlgtiyssslfpnrapsgrvlllnyiggatntgilsktegelveavdrdl 453

Qy      463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
            ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db      454 rkmlinpnakdplvlgvrvwpkaipqflvghldlldsakmalrdsgfhglflggnyvsgv 513

Qy      523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
            |||||||||||||:||  |||:||||
Db      514 algrcvegayevaaevkeflsqyayk 539


RESULT   4
AAW41609
ID   AAW41609 standard; Protein; 539 AA.
XX
AC   AAW41609;
XX
DT   20-APR-1998  (first entry)
XX
DE   Cotton protox-1.
XX
KW   Protoporphyrinogen oxidase-1; protox-1; promoter; cotton;
KW   herbicide resistance; breeding programme; probe; gene isolation;
KW   genomic mapping.
XX
OS   Gossypium hirsutum.
XX
PN   WO9732028-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US03343.
XX
PR   21-JUN-1996;   96US-0020003.
PR   28-FEB-1996;   96US-0012705.
PR   28-FEB-1996;   96US-0013612.
XX
PA   (NOVS ) NOVARTIS AG.
XX
PI   Johnson MA,  Volrath SL,  Ward ER;
XX
DR   WPI; 1997-489209/45.
DR   N-PSDB; AAV04313.
XX
PT   DNA containing a plant proto-porphyrinogen oxidase gene promoter -
PT   optionally linked to a heterologous gene, especially to express
PT   herbicide-resistant enzymes, and plants containing such constructs
XX
PS   Claim 35; Pages 80-82; 114pp; English.
XX
CC   The present sequence is cotton protoporphyrinogen oxidase-1
CC   (protox-1).
CC   The protox-1 promoter can be used to express herbicide resistant
CC   enzymes, specifically protox, i.e. a plant tissue, plant or progeny
CC   containing a chimeric gene of the promoter and a heterologous
CC   coding sequence. The plant can also be used in breeding programmes.
CC   Also hybridising fragments of the protox coding sequence can be
CC   used as probes, e.g. to isolate related genes or for genomic
CC   mapping.
XX
SQ   Sequence   539 AA;

  Query Match             74.3%;  Score 2113;  DB 18;  Length 539;
  Best Local Similarity   79.2%;  Pred. No. 2.8e-194;
```

```
  Matches  401;  Conservative   46;  Mismatches   55;  Indels    4;  Gaps    1;

Qy      47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
           :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db      34 fklrcslaegptissskidggessiadcvivgggisglciaqalatkhrdvasnvivtea 93

Qy     103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
           ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db      94 rdrvggnittverdgylweegpnsfqpsdpiltmavdsglkddlvlgdpnaprfvlwegk 153

Qy     163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
           ||||||| ||||||||||| |||||||| ||:|| |||:|||||:||||||| ||||| |
Db     154 lrpvpskptdlpffdlmsiagklragfgaigirppppgyeesveefvrrnlgaevferfi 213

Qy     223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
           |||||||| |||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db     214 epfcsgvyagdpsklsmkaafgrvwkleeiggsiiggtfktiqernktpkpprdprlpkp 273

Qy     283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
           |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
Db     274 kgqtvgsfrkgltmlpeaianslgsnvklswklssitklgnggynltfetpegmvslqsr 333

Qy     343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
           |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db     334 svvmtipshvasnllhplsaaaadalsqfyyppvasvtvsypkeairkeclidgelkgfg 393

Qy     403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
           |||||:||:||||||||||||||||| |||||||||||| |  |||||| :||| |||||
Db     394 qlhprsqgietlgtiyssslfpnrapsgrvlllnyiggatntgilsktegelveavdrdl 453

Qy     463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
           ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db     454 rkmlinpnakdplvlgvrvwpkaipqflvghldlldsakmalrdsgfhglflggnyvsgv 513

Qy     523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
           |||||||||||||:||  |||:||||
Db     514 algrcvegayevaaevkeflsqyayk 539


RESULT   5
AAW25740
ID   AAW25740 standard; Protein; 539 AA.
XX
AC   AAW25740;
XX
DT   01-MAR-1998  (first entry)
XX
DE   Cotton protoporphyrinogen oxidase (protox-1).
XX
KW   Protox-1; protoporphyrinogen oxidase; inhibitor; cotton;
KW   herbicide tolerance; herbicide resistance; transgenic plant.
XX
OS   Gossypium hirsutum L.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 365
FT                   /note= "substitution of Pro-365 by another amino
FT                          acid, especially Ser, provides a modified
FT                          protox tolerant to a herbicide (Claims
FT                          51-52)"
FT   Misc-difference 428
FT                   /note= "substitution of Tyr-428 by another amino
FT                          acid, especially Cys or Arg, provides a
FT                          modified protox tolerant to a herbicide
FT                          (Claims 53-54)"
XX
PN   WO9732011-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US03313.
XX
PR   21-JUN-1996;   96US-0020003.
PR   28-FEB-1996;   96US-0012705.
PR   28-FEB-1996;   96US-0013612.
XX
PA   (NOVS ) NOVARTIS AG.
XX
PI   Heifetz PB,  Johnson MA,  Potter SL,  Volrath SL,  Ward ER;
XX
DR   WPI; 1997-448683/41.
DR   N-PSDB; AAT86123.
XX
PT   New DNA encoding plant protoporphyrinogen oxidase enzyme  - and
PT   herbicide resistant mutants, useful to prepare plants resistant to
PT   herbicide which therefore kills undesired vegetation only
XX
PS   Claim 7; Page 139-141; 196pp; English.
XX
CC   This protein comprises cotton protoporphyrinogen oxidase
CC   (protox-1), an enzyme that catalyses the oxidation of
CC   protoporphyrinogen IX to protoporphyrin IX.  Its amino acid
CC   sequence was deduced from an cotyledon Protox-1 cDNA clone (see
CC   AAT86123).  Sites within the cotton Protox-1 gene have been
CC   identified that can be mutated to encode a modified protox that is
CC   resistant to protox inhibitors and hence tolerant of certain
CC   herbicides.  Plants, especially crop plants, may be engineered for
CC   resistance to protox inhibitors via mutation of the native protox
CC   gene to a resistant form, or they may be transformed with a gene
CC   encoding an inhibitor-resistant form of a plant protox enzyme, such
CC   as claimed forms from wheat, soybean, cotton, sugarbeet, oilseed
CC   rape, rice and sorghum (see AAW25738-48).  Application of herbicide
CC   will then kill undesired vegetation only.  Protox enzymes can also
CC   be expressed in transformed host cells and used to identify
CC   inhibitors of protox enzyme activity, i.e. herbicide candidates, or
CC   to design herbicide tolerant forms of the enzyme.
XX
SQ   Sequence   539 AA;


  Query Match             74.3%;  Score 2113;  DB 18;  Length 539;
  Best Local Similarity   79.2%;  Pred. No. 2.8e-194;
  Matches  401;  Conservative   46;  Mismatches   55;  Indels    4;  Gaps    1;

Qy      47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
           :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db      34 fklrcslaegptissskidggessiadcvivgggisglciaqalatkhrdvasnvivtea 93

Qy     103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
           ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db      94 rdrvggnittverdgylweegpnsfqpsdpiltmavdsglkddlvlgdpnaprfvlwegk 153

Qy     163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
           ||||||| ||||||||||| |||||||| ||:|| |||:|||||:||||||| ||||| |
Db     154 lrpvpskptdlpffdlmsiagklragfgaigirppppgyeesveefvrrnlgaevferfi 213

Qy     223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
           |||||||| |||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db     214 epfcsgvyagdpsklsmkaafgrvwkleeiggsiiggtfktiqernktpkpprdprlpkp 273

Qy     283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
           |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
Db     274 kgqtvgsfrkgltmlpeaianslgsnvklswklssitklgnggynltfetpegmvslqsr 333

Qy     343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
           |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db     334 svvmtipshvasnllhplsaaaadalsqfyyppvasvtvsypkeairkeclidgelkgfg 393

Qy     403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
           |||||:||:||||||||||||||||| |||||||||||| |  |||||| :||| |||||
Db     394 qlhprsqgietlgtiyssslfpnrapsgrvlllnyiggatntgilsktegelveavdrdl 453

Qy     463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
           ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db     454 rkmlinpnakdplvlgvrvwpkaipqflvghldlldsakmalrdsgfhglflggnyvsgv 513
```

```
Qy    523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
          |||||||||||||:||  |||:||||
Db    514 algrcvegayevaaevkeflsqyayk 539

RESULT   6
AAB72907
ID   AAB72907 standard; Protein; 539 AA.
XX
AC   AAB72907;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Cotton protoporphyrinogen oxidase SEQ ID NO: 16.
XX
KW   Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
KW   soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein.
XX
OS   Gossypium hirsutum.
XX
PN   WO200112825-A1.
XX
PD   22-FEB-2001.
XX
PF   30-JUN-2000; 2000WO-EP06127.
XX
PR   13-AUG-1999;   99US-0373691.
XX
PA   (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX
PI   Johnson MA,  Volrath SL,  Heifetz PB,  Law MD;
XX
DR   WPI; 2001-234914/24.
DR   N-PSDB; AAF76579.
XX
PT   Plant DNA molecules encoding herbicide-tolerant forms of
PT   protoporphyrinogen oxidase which are useful for rationally designing
PT   new inhibitory herbicides and for producing herbicide-tolerant
PT   transgenic plants and seeds -
XX
PS   Claim 22; Page 184-187; 228pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
CC   sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
CC   oxidase (protox) enzyme. Examples of these mutants are shown in
CC   AAB72920-AAB72926. They are useful as they enable the production of
CC   herbicide-tolerant plants and seeds. The present sequence is a protox
CC   protein.
XX
SQ   Sequence   539 AA;

  Query Match             74.3%;  Score 2113;  DB 22;  Length 539;
  Best Local Similarity   79.2%;  Pred. No. 2.8e-194;
  Matches  401;  Conservative   46;  Mismatches   55;  Indels    4;  Gaps    1;

Qy     47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
          :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db     34 fklrcslaegptissskidggessiadcvivgggisglciaqalatkhrdvasnvivtea 93

Qy    103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
          ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db     94 rdrvggnittverdgylweegpnsfqpsdpiltmavdsglkddlvlgdpnaprfvlwegk 153

Qy    163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
          ||||||| ||||||||||| |||||||| ||:|| |||:|||||:||||||| ||||| |
Db    154 lrpvpskptdlpffdlmsiagklragfgaigirppppgyeesveefvrrnlgaevferfi 213

Qy    223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
          |||||||| |||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db    214 epfcsgvyagdpsklsmkaafgrvwkleeiggsiiggtfktiqernktpkpprdprlpkp 273
```

```
Qy    283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
          |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
Db    274 kgqtvgsfrkgltmlpeaianslgsnvklswklssitklgnggynltfetpegmvslqsr 333

Qy    343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
          |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db    334 svvmtipshvasnllhplsaaaadalsqfyyppvasvtvsypkeairkeclidgelkgfg 393

Qy    403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
          |||||:||:||||||||||||||||| |||||||||||| |  |||||| :||| |||||
Db    394 qlhprsqgietlgtiyssslfpnrapsgrvlllnyiggatntgilsktegelveavdrdl 453

Qy    463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
          ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db    454 rkmlinpnakdplvlgvrvwpkaipqflvghldlldsakmalrdsgfhglflggnyvsgv 513

Qy    523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
          |||||||||||||:||  |||:||||
Db    514 algrcvegayevaaevkeflsqyayk 539

RESULT   7
AAB72921
ID   AAB72921 standard; Protein; 539 AA.
XX
AC   AAB72921;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Cotton protoporphyrinogen oxidase Y428H mutant.
XX
KW   Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
KW   soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein.
XX
OS   Gossypium hirsutum.
OS   Synthetic.
XX
FH   Key              Location/Qualifiers
FT   Misc-difference  428
FT                    /note= "wild-type Tyr substituted by His"
XX
PN   WO200112825-A1.
XX
PD   22-FEB-2001.
XX
PF   30-JUN-2000; 2000WO-EP06127.
XX
PR   13-AUG-1999;   99US-0373691.
XX
PA   (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX
PI   Johnson MA,  Volrath SL,  Heifetz PB,  Law MD;
XX
DR   WPI; 2001-234914/24.
XX
PT   Plant DNA molecules encoding herbicide-tolerant forms of
PT   protoporphyrinogen oxidase which are useful for rationally designing
PT   new inhibitory herbicides and for producing herbicide-tolerant
PT   transgenic plants and seeds -
XX
PS   Claim 22; Page -; 228pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
CC   sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
CC   oxidase (protox) enzyme. Examples of these mutants are shown in
CC   AAB72920-AAB72926. They are useful as they enable the production of
CC   herbicide-tolerant plants and seeds. The present sequence is a mutant
CC   protox protein.
CC   Note: The present sequence is not shown in the specification but is
CC   derived from that given in SEQ ID NO: 16 (see AAB72907).
XX
SQ   Sequence   539 AA;
```

```
  Query Match             74.1%;  Score 2108;  DB 22;  Length 539;
  Best Local Similarity   79.1%;  Pred. No. 8.5e-194;
  Matches  400;  Conservative   47;  Mismatches   55;  Indels    4;  Gaps    1;

Qy      47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
           :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db      34 fklrcslaegptissskidggessiadcvivgggisglciaqalatkhrdvasnvivtea 93

Qy     103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
           ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db      94 rdrvggnittverdgylweegpnsfqpsdpiltmavdsglkddlvlgdpnaprfvlwegk 153

Qy     163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
           ||||||| ||||||||||| |||||||| ||:|| |||:|||||:||||||| ||||| |
Db     154 lrpvpskptdlpffdlmsiagklragfgaigirppppgyeesveefvrrnlgaevferfi 213

Qy     223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
           |||||||| |||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db     214 epfcsgvyagdpsklsmkaafgrvwkleeiggsiiggtfktiqernktpkpprdprlpkp 273

Qy     283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
           |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
Db     274 kgqtvgsfrkgltmlpeaianslgsnvklswklssitklgnggynltfetpegmvslqsr 333

Qy     343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
           |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db     334 svvmtipshvasnllhplsaaaadalsqfyyppvasvtvsypkeairkeclidgelkgfg 393

Qy     403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
           |||||:||:||||||||||||||||| |||||||:|||| |  |||||| :||| |||||
Db     394 qlhprsqgietlgtiyssslfpnrapsgrvlllnhiggatntgilsktegelveavdrdl 453

Qy     463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
           ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db     454 rkmlinpnakdplvlgvrvwpkaipqflvghldlldsakmalrdsgfhglflggnyvsgv 513

Qy     523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
           |||||||||||||:||  |||:||||
Db     514 algrcvegayevaaevkeflsqyayk 539

RESULT    8
AAB72922
ID   AAB72922 standard; Protein; 539 AA.
XX
AC   AAB72922;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Cotton protoporphyrinogen oxidase Y428A mutant.
XX
KW   Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
KW   soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein.
XX
OS   Gossypium hirsutum.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 428
FT                   /note= "wild-type Tyr substituted by Ala"
XX
PN   WO200112825-A1.
XX
PD   22-FEB-2001.
XX
PF   30-JUN-2000; 2000WO-EP06127.
XX
PR   13-AUG-1999;   99US-0373691.
XX
PA   (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX
PI   Johnson MA,  Volrath SL,  Heifetz PB,  Law MD;
XX
DR   WPI; 2001-234914/24.
XX
PT   Plant DNA molecules encoding herbicide-tolerant forms of
PT   protoporphyrinogen oxidase which are useful for rationally designing
PT   new inhibitory herbicides and for producing herbicide-tolerant
PT   transgenic plants and seeds -
XX
PS   Claim 22; Page -; 228pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
CC   sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
CC   oxidase (protox) enzyme. Examples of these mutants are shown in
CC   AAB72920-AAB72926. They are useful as they enable the production of
CC   herbicide-tolerant plants and seeds. The present sequence is a mutant
CC   protox protein.
CC   Note: The present sequence is not shown in the specification but is
CC   derived from that given in SEQ ID NO: 16 (see AAB72907).
XX
SQ   Sequence   539 AA;

  Query Match             74.0%;  Score 2104;  DB 22;  Length 539;
  Best Local Similarity   79.1%;  Pred. No. 2.1e-193;
  Matches  400;  Conservative   46;  Mismatches   56;  Indels    4;  Gaps    1;

Qy      47 WRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANY----PNLMVTEA 102
           :: |||:|:  |: || :||| ::  |||||| ||||||||| ::  :     |::||||
Db      34 fklrcslaegptissskidggessiadcvivgggisglciaqalatkhrdvasnvivtea 93

Qy     103 RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWKGK 162
           ||| ||||||||||||||||||||||||||:|||||| |||||||||||||||||||:||
Db      94 rdrvggnittverdgylweegpnsfqpsdpiltmavdsglkddlvlgdpnaprfvlwegk 153

Qy     163 LRPVPSKLTDLPFFDLMSIPGKLRAGFGPIGLRPSPPGHEESVEQFVRRNLGGEVFERLI 222
           ||||||| ||||||||||| |||||||| ||:|| |||:|||||:||||||| ||||| |
Db     154 lrpvpskptdlpffdlmsiagklragfgaigirppppgyeesveefvrrnlgaevferfi 213

Qy     223 EPFCSGVYVGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKP 282
           |||||||| |||||||||||||:|||||| |||||||||| |:||: ||| |||||||||
Db     214 epfcsgvyagdpsklsmkaafgrvwkleeiggsiiggtfktiqernktpkpprdprlpkp 273

Qy     283 KGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPEGVVSLQSR 342
           |||||||||||| |||:||:  ||| :|||||||||||   |||:||:|||||:||||||
Db     274 kgqtvgsfrkgltmlpeaianslgsnvklswklssitklgnggynltfetpegmvslqsr 333

Qy     343 SIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDERLVDGELKGFG 402
           |:|||:||:||||:| ||| ||||||| |||||| :||:|||:|||| | |:||||||||
Db     334 svvmtipshvasnllhplsaaaadalsqfyyppvasvtvsypkeairkeclidgelkgfg 393

Qy     403 QLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKTESQLVEVVDRDL 462
           |||||:||:||||||||||||||||| ||||||| |||| |  |||||| :||| |||||
Db     394 qlhprsqgietlgtiyssslfpnrapsgrvlllnaiggatntgilsktegelveavdrdl 453

Qy     463 RKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNGLEGLFLGGNYVSGV 522
           ||||| | |:||||:||||||:||||||||||| | :|| |: |:|  ||||||||||||
Db     454 rkmlinpnakdplvlgvrvwpkaipqflvghldlldsakmalrdsgfhglflggnyvsgv 513

Qy     523 ALGRCVEGAYEVASEVTGFLSRYAYK 548
           |||||||||||||:||  |||:||||
Db     514 algrcvegayevaaevkeflsqyayk 539

RESULT    9
AAB72923
ID   AAB72923 standard; Protein; 537 AA.
XX
AC   AAB72923;
XX
DT   16-MAY-2001  (first entry)
XX
```

```
DE   Arabidopsis protoporphyrinogen oxidase A220I mutant.
XX
KW   Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
KW   soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein.
XX
OS   Arabidopsis thaliana.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 220
FT                   /note= "wild-type Ala substituted by Ile"
XX
PN   WO200112825-A1.
XX
PD   22-FEB-2001.
XX
PF   30-JUN-2000; 2000WO-EP06127.
XX
PR   13-AUG-1999;   99US-0373691.
XX
PA   (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX
PI   Johnson MA,  Volrath SL,  Heifetz PB,  Law MD;
XX
DR   WPI; 2001-234914/24.
XX
PT   Plant DNA molecules encoding herbicide-tolerant forms of
PT   protoporphyrinogen oxidase which are useful for rationally designing
PT   new inhibitory herbicides and for producing herbicide-tolerant
PT   transgenic plants and seeds -
XX
PS   Claim 23; Page -; 228pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
CC   sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
CC   oxidase (protox) enzyme. Examples of these mutants are shown in
CC   AAB72920-AAB72926. They are useful as they enable the production of
CC   herbicide-tolerant plants and seeds. The present sequence is a mutant
CC   protox protein.
CC   Note: The present sequence is not shown in the specification but is
CC   derived from that given in SEQ ID NO: 2 (see AAB72901).
XX
SQ   Sequence   537 AA;

  Query Match             73.1%;  Score 2078;  DB 22;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 6.5e-191;
  Matches  404;  Conservative   50;  Mismatches   72;  Indels   10;  Gaps    4;

Qy      22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
           |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db       3 lsllrpttqsllpsfskpn-lrlnvykplrlrcsvaggptvgsskieggggttittdcvi 61

Qy      77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
           || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db      62 vgggisglciaqalatkhpdaapnlivteakdrvggniitreengflweegpnsfqpsdp 121

Qy     133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
           |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db     122 mltmvvdsglkddlvlgdptaprfvlwngklrpvpskltdlpffdlmsiggkiragfgal 181

Qy     193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
           |:|||||| |||||:||||||| |||||||||||||||:|||||||||||||||||||:
Db     182 girpsppgreesveefvrrnlgdevferliepfcsgvyigdpsklsmkaafgkvwkleqn 241

Qy     253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
           ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db     242 ggsiiggtfkaiqerknapkaerdprlpkpqgqtvgsfrkglrmlpeaisarlgskvkls 301

Qy     313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
           |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db     302 wklsgitklesggynltyetpdglvsvqsksvvmtvpshvasgllrplsesaanalskly 361
```

```
Qy     373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
           |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db     362 yppvaavsisypkeairteclidgelkgfgqlhprtqgvetlgtiyssslfpnrappgri 421

Qy     433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
           ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db     422 lllnyiggstntgilsksegelveavdrdlrkmlikpnstdplklgvrvwpqaipqflvg 481

Qy     493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
           | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db     482 hfdildtakssltssgyeglflggnyvagvalgrcvegayetaievnnfmsryayk 537

RESULT  10
AAR90295
ID   AAR90295 standard; Protein; 537 AA.
XX
AC   AAR90295;
XX
DT   02-JUL-1996  (first entry)
XX
DE   Protein having protoporphyrinogen oxidase activity.
XX
KW   Protoporphyrinogen oxidase; PPO; herbicide; resistance; mutant
KW   variegate porphyria.
XX
OS   Arabidopsis sp.
XX
PN   WO9534659-A1.
XX
PD   21-DEC-1995.
XX
PF   08-JUN-1995;   95WO-IB00452.
XX
PR   16-JUN-1994;   94US-0261198.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Volrath S,  Ward ER;
XX
DR   WPI; 1996-049687/05.
DR   N-PSDB; AAT11673.
XX
PT   DNA encoding eukaryotic proto:porphyrinogen oxidase and herbicide
PT   resistant mutants  - used to make herbicide resistant plants and for
PT   diagnosis and treatment of variegate porphyria
XX
PS   Claim 6; Page 71-73; 118pp; English.
XX
CC   DNA encoding a modified protoporphyrinogen oxidase (PPO) or a
CC   chimeric gene comprising a promoter (pref. active in a plant) linked
CC   to such a DNA or DNA encoding a wild type PPO can be used (1) to
CC   impart herbicide resistance to plants; (2) for treating and
CC   diagnosing deficient PPO activity in animals (esp. variegate
CC   porphyria); and (3) for the production of recombinant PPO which is
CC   useful as an assay reagent and in rational design of new inhibitory
CC   herbicides. Herbicide resistant PPO genes can also be used to
CC   select plants transformed with a transgene and probes derived from
CC   the genes can be used to quantify levels of PPO mRNA.
XX
SQ   Sequence   537 AA;

  Query Match             73.0%;  Score 2075;  DB 17;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 1.3e-190;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy      22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
           |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db       3 lsllrpttqsllpsfskpn-lrlnvykplrlrcsvaggptvgsskieggggttittdcvi 61

Qy      77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
```

```
          || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db     62 vgggisglciaqalatkhpdaapnlivteakdrvggniitreengflweegpnsfqpsdp 121

Qy    133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
          |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db    122 mltmvvdsglkddlvlgdptaprfvlwngklrpvpskltdlpffdlmsiggkiragfgal 181

Qy    193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
          |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||: 
Db    182 girpsppgreesveefvrrnlgdevferliepfcsgvyagdpsklsmkaafgkvwkleqn 241

Qy    253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
          ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db    242 ggsiiggtfkaiqerknapkaerdprlpkpqgqtvgsfrkglrmlpeaisarlgskvkls 301

Qy    313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
          |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db    302 wklsgitklesggynltyetpdglvsvqsksvvmtvpshvasgllrplsesaanalskly 361

Qy    373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
          |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db    362 yppvaavsisypkeairteclidgelkgfgqlhprtqgvetlgtiyssslfpnrappgri 421

Qy    433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
          ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db    422 lllnyiggstntgilsksegelveavdrdlrkmlikpnstdplklgvrvwpqaipqflvg 481

Qy    493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
          | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db    482 hfdildtaksslssgyeglflggnyvagvalgrcvegayetaievnnfmsryayk 537

RESULT  11
AAW41603
ID   AAW41603 standard; Protein; 537 AA.
XX
AC   AAW41603;
XX
DT   20-APR-1998  (first entry)
XX
DE   Arabidopsis thaliana protox-1.
XX
KW   Protoporphyrinogen oxidase-1; protox-1; promoter;
KW   herbicide resistance; breeding programme; probe; gene isolation;
KW   genomic mapping.
XX
OS   Arabidopsis thaliana.
XX
PN   WO9732028-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US03343.
XX
PR   21-JUN-1996;   96US-0020003.
PR   28-FEB-1996;   96US-0012705.
PR   28-FEB-1996;   96US-0013612.
XX
PA   (NOVS ) NOVARTIS AG.
XX
PI   Johnson MA,  Volrath SL,  Ward ER;
XX
DR   WPI; 1997-489209/45.
DR   N-PSDB; AAV04304.
XX
PT   DNA containing a plant proto-porphyrinogen oxidase gene promoter -
PT   optionally linked to a heterologous gene, especially to express
PT   herbicide-resistant enzymes, and plants containing such constructs
XX
PS   Claim 27; Pages 36-39; 114pp; English.
XX
CC   The present sequence is Arabidopsis thaliana
CC   protoporphyrinogen oxidase-1 (protox-1).
```

```
CC   The protox-1 promoter can be used to express herbicide resistant
CC   enzymes, specifically protox, i.e. a plant tissue, plant or progeny
CC   containing a chimeric gene of the promoter and a heterologous
CC   coding sequence. The plant can also be used in breeding programmes.
CC   Also hybridising fragments of the protox coding sequence can be
CC   used as probes, e.g. to isolate related genes or for genomic
CC   mapping.
XX
SQ   Sequence   537 AA;

  Query Match             73.0%;  Score 2075;  DB 18;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 1.3e-190;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy     22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
          |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db      3 lsllrpttqsllpsfskpn-lrlnvykplrlrcsvaggptvgsskieggggttittdcvi 61

Qy     77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
          || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db     62 vgggisglciaqalatkhpdaapnlivteakdrvggniitreengflweegpnsfqpsdp 121

Qy    133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
          |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db    122 mltmvvdsglkddlvlgdptaprfvlwngklrpvpskltdlpffdlmsiggkiragfgal 181

Qy    193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
          |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||: 
Db    182 girpsppgreesveefvrrnlgdevferliepfcsgvyagdpsklsmkaafgkvwkleqn 241

Qy    253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
          ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db    242 ggsiiggtfkaiqerknapkaerdprlpkpqgqtvgsfrkglrmlpeaisarlgskvkls 301

Qy    313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
          |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db    302 wklsgitklesggynltyetpdglvsvqsksvvmtvpshvasgllrplsesaanalskly 361

Qy    373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
          |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db    362 yppvaavsisypkeairteclidgelkgfgqlhprtqgvetlgtiyssslfpnrappgri 421

Qy    433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
          ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db    422 lllnyiggstntgilsksegelveavdrdlrkmlikpnstdplklgvrvwpqaipqflvg 481

Qy    493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
          | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db    482 hfdildtaksslssgyeglflggnyvagvalgrcvegayetaievnnfmsryayk 537

RESULT  12
AAW25746
ID   AAW25746 standard; Protein; 537 AA.
XX
AC   AAW25746;
XX
DT   01-MAR-1998  (first entry)
XX
DE   Arabidopsis protoporphyrinogen oxidase (protox-1).
XX
KW   Protox-1; protoporphyrinogen oxidase; inhibitor;
KW   herbicide tolerance; herbicide resistance; transgenic plant.
XX
OS   Arabidopsis thaliana.
XX
PN   WO9732011-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US03313.
XX
```

```
PR   21-JUN-1996;   96US-0020003.
PR   28-FEB-1996;   96US-0012705.
PR   28-FEB-1996;   96US-0013612.
XX
PA   (NOVS ) NOVARTIS AG.
XX
PI   Heifetz PB,  Johnson MA,  Potter SL,  Volrath SL,  Ward ER;
XX
DR   WPI; 1997-448683/41.
DR   N-PSDB; AAT86129.
XX
PT   New DNA encoding plant protoporphyrinogen oxidase enzyme  - and
PT   herbicide resistant mutants, useful to prepare plants resistant to
PT   herbicide which therefore kills undesired vegetation only
XX
PS   Claim 95; Page 95-98; 196pp; English.
XX
CC   This protein comprises Arabidopsis protoporphyrinogen oxidase
CC   (protox-1), an enzyme that catalyses the oxidation of
CC   protoporphyrinogen IX to protoporphyrin IX.  Its amino acid sequence
CC   was deduced from a previously obtained Protox-1 cDNA clone (see
CC   AAT86129).  Arabidopsis protox-1 can be modified to render it
CC   resistant to protox inhibitors and hence tolerant of certain
CC   herbicides.  Plants, especially crop plants, may be engineered for
CC   resistance to protox inhibitors via mutation of the native protox
CC   gene to a resistant form, or they may be transformed with a gene
CC   encoding an inhibitor-resistant form of a plant protox enzyme, such
CC   as claimed forms from wheat, soybean, cotton, sugarbeet, oilseed
CC   rape, rice and sorghum (see AAW25738-48).  Application of herbicide
CC   will then kill undesired vegetation only.  Protox enzymes can also
CC   be expressed in transformed host cells and used to identify
CC   inhibitors of protox enzyme activity, i.e. herbicide candidates, or
CC   to design herbicide tolerant forms of the enzyme.
XX
SQ   Sequence   537 AA;


  Query Match             73.0%;  Score 2075;  DB 18;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 1.3e-190;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy         22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
              |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db          3 lsllrpttqsllpsfskpn-lrlnvykplrlrcsvaggptvgsskieggggttittdcvi 61

Qy         77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
              || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db         62 vgggisglciaqalatkhpdaapnlivteakdrvggniitreengflweegpnsfqpsdp 121

Qy        133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
              |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db        122 mltmvvdsglkddlvlgdptaprfvlwngklrpvpskltdlpffdlmsiggkiragfgal 181

Qy        193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
              |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||:
Db        182 girpsppgreesveefvrrnlgdevferliepfcsgvyagdpsklsmkaafgkvwkleqn 241

Qy        253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
              ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db        242 ggsiiggtfkaiqerknapkaerdprlpkpqgqtvgsfrkglrmlpeaisarlgskvkls 301

Qy        313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
              |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db        302 wklsgitklesggynltyetpdglvsvqsksvvmtvpshvasgllrplsesaanalskly 361

Qy        373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
              |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db        362 yppvaavsisypkeairteclidgelkgfgqlhprtqgvetlgtiyssslfpnrappgri 421

Qy        433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
              ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db        422 lllnyiggstntgilsksegelveavdrdlrkmlikpnstdplklgvrvwpqaipqflvg 481
```

```
Qy        493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
              | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db        482 hfdildtaksslts sgyeglflggnyvagvalgrcvegayetaievnnfmsryayk 537

RESULT  13
AAW51347
ID   AAW51347 standard; Protein; 537 AA.
XX
AC   AAW51347;
XX
DT   14-AUG-1998  (first entry)
XX
DE   Arabidopsis proto-porphyrinogen oxidase-1 (protox-1).
XX
KW   Proto-porphyrinogen oxidase; protox; herbicide tolerant; resistance;
KW   inhibitor; variegate porphyria.
XX
OS   Arabidopsis thaliana.
XX
PN   US5767373-A.
XX
PD   16-JUN-1998.
XX
PF   16-JUN-1994;   94US-0261198.
XX
PR   06-JUN-1995;   95US-0472028.
PR   16-JUN-1994;   94US-0261198.
XX
PA   (NOVS ) NOVARTIS FINANCE CORP.
XX
PI   Volrath S,  Ward ER;
XX
DR   WPI; 1998-361821/31.
DR   N-PSDB; AAV07251.
XX
PT   DNA encoding eukaryote herbicide resistant proto-porphyrinogen
PT   oxidase - useful for producing recombinant plants having functional
PT   enzyme, to be grown in the presence of herbicides
XX
PS   Claim 6; Columns 21-22; 43pp; English.
XX
CC   The invention relates to eukaryotic DNA sequences coding for native
CC   proto-porphyrinogen oxidase (protox) or modified forms of the enzyme
CC   which are herbicide tolerant. Plants having altered protox activity
CC   which confers tolerance to herbicides are also provided. These plants
CC   may be bred or engineered for resistance to protox inhibitors via
CC   mutation of the native protox gene to a resistant form or through
CC   increased levels of expression of the native protox gene, or they may be
CC   transformed with modified eukaryotic or prokaryotic protox coding
CC   sequences or wild type prokaryotic protox sequences which are herbicide
CC   tolerant. In the human condition variegate porphyria, an autosomal
CC   dominant disorder characterised by neuropsychiatric symptoms, decreased
CC   protox activity is detected. Thus the DNA molecules can be used to
CC   provide probes to detect and quantify protox levels in the diagnosis of
CC   diseases associated with the enzyme. The present sequence represents
CC   Arabidopsis proto-porphyrinogen oxidase-1 (protox-1).
XX
SQ   Sequence   537 AA;


  Query Match             73.0%;  Score 2075;  DB 19;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 1.3e-190;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy         22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
              |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db          3 lsllrpttqsllpsfskpn-lrlnvykplrlrcsvaggptvgsskieggggttittdcvi 61

Qy         77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
              || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db         62 vgggisglciaqalatkhpdaapnlivteakdrvggniitreengflweegpnsfqpsdp 121
```

```
Qy     133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
           |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db     122 mltmvvdsglkddlvlgdptaprfvlwngklrpvpskltdlpffdlmsiggkiragfgal 181

Qy     193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
           |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||:
Db     182 girpsppgreesveefvrrnlgdevferliepfcsgvyagdpsklsmkaafgkvwkleqn 241

Qy     253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
           ||||||||||||:|| :  ||| ||||||||:|||||||||||||||:|||||||||:|||
Db     242 ggsiiggtfkaiqerknapkaerdprlpkpqgqtvgsfrkglrmlpeaisarlgskvkls 301

Qy     313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
           |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db     302 wklsgitklesggynltyetpdglvsvqsksvvmtvpshvasgllrplsesaanalskly 361

Qy     373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
           |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db     362 yppvaavsisypkeairteclidgelkgfgqlhprtqgvetlgtiyssslfpnrappgri 421

Qy     433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
           ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db     422 lllnyiggstntgilsksegelveavdrdlrkmlikpnstdplklgvrvwpqaipqflvg 481

Qy     493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
           | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db     482 hfdildtakssltssgyeglflggnyvagvalgrcvegayetaievnnfmsryayk 537


RESULT  14
AAG42385
ID   AAG42385 standard; Protein; 537 AA.
XX
AC   AAG42385;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 52855.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
```

```
PR   23-JUL-1999;    99US-0145218.
PR   23-JUL-1999;    99US-0145224.
PR   26-JUL-1999;    99US-0145276.
PR   27-JUL-1999;    99US-0145913.
PR   27-JUL-1999;    99US-0145918.
PR   27-JUL-1999;    99US-0145919.
PR   28-JUL-1999;    99US-0145951.
PR   02-AUG-1999;    99US-0146386.
PR   02-AUG-1999;    99US-0146388.
PR   02-AUG-1999;    99US-0146389.
PR   03-AUG-1999;    99US-0147038.
PR   04-AUG-1999;    99US-0147204.
PR   04-AUG-1999;    99US-0147302.
PR   05-AUG-1999;    99US-0147192.
PR   05-AUG-1999;    99US-0147260.
PR   06-AUG-1999;    99US-0147303.
PR   06-AUG-1999;    99US-0147416.
PR   09-AUG-1999;    99US-0147493.
PR   09-AUG-1999;    99US-0147935.
PR   10-AUG-1999;    99US-0148171.
PR   11-AUG-1999;    99US-0148319.
PR   12-AUG-1999;    99US-0148341.
PR   13-AUG-1999;    99US-0148565.
PR   13-AUG-1999;    99US-0148684.
PR   16-AUG-1999;    99US-0149368.
PR   17-AUG-1999;    99US-0149175.
PR   18-AUG-1999;    99US-0149426.
PR   20-AUG-1999;    99US-0149722.
PR   20-AUG-1999;    99US-0149723.
PR   20-AUG-1999;    99US-0149929.
PR   23-AUG-1999;    99US-0149902.
PR   23-AUG-1999;    99US-0149930.
PR   25-AUG-1999;    99US-0150566.
PR   26-AUG-1999;    99US-0150884.
PR   27-AUG-1999;    99US-0151065.
PR   27-AUG-1999;    99US-0151066.
PR   27-AUG-1999;    99US-0151080.
PR   30-AUG-1999;    99US-0151303.
PR   31-AUG-1999;    99US-0151438.
PR   01-SEP-1999;    99US-0151930.
PR   07-SEP-1999;    99US-0152363.
PR   10-SEP-1999;    99US-0153070.
PR   13-SEP-1999;    99US-0153758.
PR   15-SEP-1999;    99US-0154018.
PR   16-SEP-1999;    99US-0154039.
PR   20-SEP-1999;    99US-0154779.
PR   22-SEP-1999;    99US-0155139.
PR   23-SEP-1999;    99US-0155486.
PR   24-SEP-1999;    99US-0155659.
PR   28-SEP-1999;    99US-0156458.
PR   29-SEP-1999;    99US-0156596.
PR   04-OCT-1999;    99US-0157117.
PR   05-OCT-1999;    99US-0157753.
PR   06-OCT-1999;    99US-0157865.
PR   07-OCT-1999;    99US-0158029.
PR   08-OCT-1999;    99US-0158232.
PR   12-OCT-1999;    99US-0158369.
PR   13-OCT-1999;    99US-0159293.
PR   13-OCT-1999;    99US-0159294.
PR   13-OCT-1999;    99US-0159295.
PR   14-OCT-1999;    99US-0159329.
PR   14-OCT-1999;    99US-0159330.
PR   14-OCT-1999;    99US-0159331.
PR   14-OCT-1999;    99US-0159637.
PR   14-OCT-1999;    99US-0159638.
PR   18-OCT-1999;    99US-0159584.
PR   21-OCT-1999;    99US-0160741.
PR   21-OCT-1999;    99US-0160767.
PR   21-OCT-1999;    99US-0160768.
PR   21-OCT-1999;    99US-0160770.
PR   21-OCT-1999;    99US-0160814.
PR   21-OCT-1999;    99US-0160815.
PR   22-OCT-1999;    99US-0160980.
PR   22-OCT-1999;    99US-0160981.
PR   22-OCT-1999;    99US-0160989.
PR   25-OCT-1999;    99US-0161404.
PR   25-OCT-1999;    99US-0161405.
PR   25-OCT-1999;    99US-0161406.
PR   26-OCT-1999;    99US-0161359.
PR   26-OCT-1999;    99US-0161360.
PR   26-OCT-1999;    99US-0161361.
PR   28-OCT-1999;    99US-0161920.
PR   28-OCT-1999;    99US-0161992.
PR   28-OCT-1999;    99US-0161993.
PR   29-OCT-1999;    99US-0162142.

  Query Match             73.0%;  Score 2075;  DB 21;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 1.3e-190;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy     22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
          |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db      3 lsllrpttqsllpsfskpn-lrlnvykplrlrcsvaggptvgsskieggggttittdcvi 61

Qy     77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
          || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db     62 vgggisglciaqalatkhpdaapnlivteakdrvggniitreengflweegpnsfqpsdp 121

Qy    133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
          |||| || ||||||||||| ||||||| ||||||||||||||||||||| ||:||||| :
Db    122 mltmvvdsglkddlvlgdptaprfvlwngklrpvpskltdlpffdlmsiggkiragfgal 181

Qy    193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
          |:|||||| |||||:||||||| ||||||||||||||| |||||||||||||||||||: 
Db    182 girpsppgreesveefvrrnlgdevferliepfcsgvyagdpsklsmkaafgkvwkleqn 241

Qy    253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
          ||||||||||||:|| : ||| ||||||||:|||||||||||||||:|||||||||:|||
Db    242 ggsiiggtfkaiqerknapkaerdprlpkpqgqtvgsfrkglrmlpeaisarlgskvkls 301

Qy    313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
          |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db    302 wklsgitklesggynltyetpdglvsvqsksvvmtvpshvasgllrplsesaanalskly 361

Qy    373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
          |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db    362 yppvaavsisypkeairteclidgelkgfgqlhprtqgvetlgtiysssllfpnrappgri 421

Qy    433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
          ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db    422 lllnyiggstntgilsksegelveavdrdlrkmlikpnstdplklgvrvwpqaipqflvg 481

Qy    493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
          | | | |||:::  :| ||||||||||:||||||||||||| | ||  |:||||||
Db    482 hfdildtaksssltssgyeglflggnyvagvalgrcvegayetaievnnfmsryayk 537

RESULT  15
AAB72901
ID   AAB72901 standard; Protein; 537 AA.
XX
AC   AAB72901;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Arabidopsis protoporphyrinogen oxidase SEQ ID NO: 2.
XX
KW   Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
KW   soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200112825-A1.
XX
PD   22-FEB-2001.
XX
```

```
PF   30-JUN-2000; 2000WO-EP06127.
XX
PR   13-AUG-1999;   99US-0373691.
XX
PA   (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX
PI   Johnson MA,  Volrath SL,  Heifetz PB,  Law MD;
XX
DR   WPI; 2001-234914/24.
DR   N-PSDB; AAF76571.
XX
PT   Plant DNA molecules encoding herbicide-tolerant forms of
PT   protoporphyrinogen oxidase which are useful for rationally designing
PT   new inhibitory herbicides and for producing herbicide-tolerant
PT   transgenic plants and seeds -
XX
PS   Claim 23; Page 137-140; 228pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
CC   sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
CC   oxidase (protox) enzyme. Examples of these mutants are shown in
CC   AAB72920-AAB72926. They are useful as they enable the production of
CC   herbicide-tolerant plants and seeds. The present sequence is a protox
CC   protein.
XX
SQ   Sequence   537 AA;


  Query Match             73.0%;  Score 2075;  DB 22;  Length 537;
  Best Local Similarity   75.4%;  Pred. No. 1.3e-190;
  Matches  404;  Conservative   49;  Mismatches   73;  Indels   10;  Gaps    4;

Qy         22 LAFLNRTSFIPFSSISKRNSVNCNGW---RTRCSVAKDYTVPSSAVDGGPAAEL--DCVI 76
              |: |  |:     | || | :  | :   | |||||   || || ::||    :  ||||
Db          3 lsllrpttqsllpsfskpn-lrlnvykplrlrcsvaggptvgsskieggggttittdcvi 61

Qy         77 VGAGISGLCIAQVMSANY----PNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDP 132
              || ||||||||| ::  :    |||:||||:|| |||| | | :|:||||||||||||||
Db         62 vgggisglciaqalatkhpdaapnlivteakdrvggniitreengflweegpnsfqpsdp 121

Qy        133 MLTMAVDCGLKDDLVLGDPNAPRFVLWKGKLRPVPSKLTDLPFFDLMSIPGKLRAGFGPI 192
              |||| || ||||||||||| ||||||| |||||||||||||||||||||| ||:||||| :
Db        122 mltmvvdsglkddlvlgdptaprfvlwngklrpvpskltdlpffdlmsiggkiragfgal 181

Qy        193 GLRPSPPGHEESVEQFVRRNLGGEVFERLIEPFCSGVYVGDPSKLSMKAAFGKVWKLEET 252
              |:|||||| |||||:||||||| |||||||||||||||| |||||||||||||||||||: 
Db        182 girpsppgreesveefvrrnlgdevferliepfcsgvyagdpsklsmkaafgkvwkleqn 241

Qy        253 GGSIIGGTFKAIKERSSTPKAPRDPRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLS 312
              ||||||||||||:|| : ||| ||||||||:|||||||||||||||:||||||||||:|||
Db        242 ggsiiggtfkaiqerknapkaerdprlpkpqgqtvgsfrkglrmlpeaisarlgskvkls 301

Qy        313 WKLSSITKSEKGGYHLTYETPEGVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFY 372
              |||| ||| | |||:||||||:|:||:||:|:||||||:||| :||||| :||:|||  |
Db        302 wklsgitklesggynltyetpdglvsvqsksvvmtvpshvasgllrplsesaanalskly 361

Qy        373 YPPVGAVTISYPQEAIRDERLVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRV 432
              |||| ||:||||:|||| | |:|||||||||||||||||||||||||||||||||| ||:
Db        362 yppvaavsisypkeairteclidgelkgfgqlhprtqgvetlgtiyssslfpnrappgri 421

Qy        433 LLLNYIGGAKNPEILSKTESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVG 492
              ||||||||: |  ||||:| :||| |||||||||||| : ||| :|||||||||||||||
Db        422 lllnyiggstntgilsksegelveavdrdlrkmlikpnstdplklgvrvwpqaipqflvg 481

Qy        493 HLDTLSTAKAAMNDNGLEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK 548
              | | | |||:::  :| |||||||||||:||||||||||||| | ||  |:||||||
Db        482 hfdildtaksslteggyeglflggnyvagvalgrcvegayetaievnnfmsryayk 537


Search completed: July  3, 2001, 10:32:56
Job time: 1975 sec
```

```
OM of: US-09-508-418-2 to: GenEmbl:*   out_format : pfs

Date: Jul 3, 2001 11:36 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09508418/runat_03072001_100017_24438/app_query.fasta_1.613
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09508418_@CGN1_1_5606 -NCPU=6
 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-508-418-2
Query length: 548
Database: GenEmbl:*
Database sequences: 1344157
Database length: -856060004
Search time (sec): 1837.380000


score_list:
Sequence            Strd Orig     ZScore    EScore Len   ! Documentation  ..
gb_pl1:AF044128       + 2832.00 3442.64  1.6e-183  1874   ! AF044128 Nicotiana tabacum pr
gb_pl1:AB020501       + 2824.00 3434.03  4.8e-183  1647   ! AB020501 Nicotiana tabacum mRN
gb_pl3:NTY13465       + 2821.00 3429.16  9.0e-183  1892   ! Y13465 Nicotiana tabacum mRNA
gb_pl4:STA225107      + 2565.50 3117.46  2.1e-165  1976   ! AJ225107 Solanum tuberosum (cu
gb_pat2:AX084746      + 2113.00 2566.79  9.8e-135  1826   ! AX084746 Sequence 15 from Pate
gb_pat1:A78957        + 2077.00 2523.53  2.5e-132  1704   ! A78957 Sequence 11 from Patent
gb_pat2:AX084732      + 2077.00 2523.46  2.5e-132  1719   ! AX084732 Sequence 1 from Paten
gb_pl2:ATHPPOX        + 2075.00 2521.57  3.2e-132  1614   ! D83139 Arabidopsis thaliana mR
gb_pat2:AX084742      + 2062.50 2505.16  2.7e-131  1847   ! AX084742 Sequence 11 from Pate
gb_pat2:AX084750      + 2006.50 2437.23  1.6e-127  1784   ! AX084750 Sequence 19 from Pate
gb_pl2:AF160961       + 2004.50 2434.12  2.4e-127  1925   ! AF160961 Cichorium intybus pr
gb_pl2:AF218052       + 1989.50 2417.42  2.0e-126  1608   ! AF218052 Zea mays protoporphyr
gb_pat2:AX084740      + 1987.00 2413.34  3.5e-126  1811   ! AX084740 Sequence 9 from Paten
gb_pat2:AX084736      + 1978.00 2402.97  1.3e-125  1691   ! AX084736 Sequence 5 from Paten
gb_pat1:A78959        + 1978.00 2402.93  1.3e-125  1698   ! A78959 Sequence 13 from Patent
gb_pl1:AB057749       + 1974.00 2396.97  2.8e-125  1923   ! AB057749 Oryza sativa PPOX-I/
gb_pl1:AB029492       + 1966.50 2388.97  7.9e-125  1689   ! AB029492 Spinacia oleracea SO-
gb_pat2:AX084748      + 1933.50 2347.68  1.6e-122  1910   ! AX084748 Sequence 17 from Pate
gb_pat2:AX084754      + 1823.00 2214.65  4.0e-115  1590   ! AX084754 Sequence 23 from Pate
gb_pl4:T15B16         - 1726.00 2061.01  1.5e-106  90596   ! AF104919 Arabidopsis thaliana
gb_pl2:ATCHRIV4       + 1726.00 2054.12  3.5e-106  198697   ! AL161492 Arabidopsis thalian
gb_pat2:E65171        + 1530.50 1856.97  3.4e-95   1838   ! E65171 Method for evaluating a
gb_pl1:AF068635       + 1530.50 1854.35  4.7e-95   2480   ! AF068635 Chlamydomonas reinhar
gb_pl1:AB057771       + 1524.00 1839.54  3.2e-94   5441   ! AB057771 Oryza sativa PPOX-I/
gb_pl2:AP003020       + 1524.00 1809.90  1.4e-92   159749   ! AP003020 Oryza sativa genomi
gb_pat2:AX084752      + 1261.00 1532.16  4.2e-77   1224   ! AX084752 Sequence 21 from Pate
gb_pl2:AF332964       +  988.50 1195.56  2.3e-58   2060   ! AF332964 Polytomella sp. 'Prin
gb_pl1:AF030179       +  606.00  715.42  1.3e-31   10265   ! AF030179 Chlamydomonas reinha
gb_ba1:AF098938       +  541.00  653.59  3.6e-28   1416   ! AF098938 Myxococcus xanthus pr
gb_ba3:MXACSGAX       -  535.00  637.37  2.9e-27   3909   ! L27429 Myxococcus xanthus csgA
gb_ba3:MXACSGA        -  534.50  640.26  2.0e-27   2622   ! M73709 Myxococcus xanthus csgA
gb_pat2:E21704        +  522.50  629.21  8.2e-27   1745   ! E21704 Protoporphyrinogen oxid
gb_pl1:AB025102       +  522.50  629.21  8.2e-27   1745   ! AB025102 Glycine max mRNA for
gb_ba1:AE000768       -  518.00  606.19  1.6e-25   12905   ! AE000768 Aquifex aeolicus sec
gb_pl3:NTY13466       +  506.50  608.79  1.1e-25   1940   ! Y13466 Nicotiana tabacum mRNA
gb_pl1:AB020500       +  505.00  609.13  1.1e-25   1515   ! AB020500 Nicotiana tabacum mRN
gb_pl1:AF044129       +  505.00  607.23  1.4e-25   1882   ! AF044129 Nicotiana tabacum pr
gb_pat2:E14138        +  490.50  590.65  1.2e-24   1662   ! E14138 cDNA encoding Arabidops
gb_pat2:AX084734      +  490.50  590.26  1.2e-24   1738   ! AX084734 Sequence 3 from Paten
gb_ba2:BACHEMEHY      +  485.00  571.82  1.3e-23   6624   ! M97208 Bacillus subtilis penic
gb_ba2:BSY14083       +  485.00  559.83  6.0e-23   26007   ! Y14083 Bacillus subtilis chr
```

```
gb_ba2:BSUB0006       +  485.00  541.50  6.3e-22   210440  ! Z99109 Bacillus subtilis
gb_pl2:AF273767       +  467.50  560.82  5.3e-23   2042   ! AF273767 Zea mays protoporp
gb_pat2:AX084738      +  467.50  560.74  5.4e-23   2061   ! AX084738 Sequence 7 from Pa
gb_pl4:STA225108      +  424.00  512.04  2.8e-20   1261   ! AJ225108 Solanum tuberosum

seq_name: gb_pl1:AF044128

seq_documentation_block:
LOCUS       AF044128      1874 bp    mRNA            PLN       05-JAN-1999
DEFINITION  Nicotiana tabacum protoporphyrinogen oxidase PX-1 mRNA, complete
            cds.
ACCESSION   AF044128
VERSION     AF044128.1  GI:4105185
KEYWORDS    .
SOURCE      common tobacco.
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1874)
  AUTHORS   Horikoshi,M., Mametsuka,K. and Hirooka,T.
  TITLE     The Molecular Basis of Photobleaching Herbicide Resistance in
            Tobacco
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1874)
  AUTHORS   Horikoshi,M., Mametsuka,K. and Hirooka,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JAN-1998) Research Center, Nihon Nohyaku Co. Ltd.,
            345, Oyamada-cho, Kawachi-nagano, Osaka 586-0094, Japan
FEATURES             Location/Qualifiers
     source          1. .1874
                     /organism="Nicotiana tabacum"
                     /cultivar="SR1"
                     /db_xref="taxon:4097"
     CDS             26. .1672
                     /EC_number="1.3.3.4"
                     /codon_start=1
                     /product="protoporphyrinogen oxidase PX-1"
                     /protein_id="AAD02290.1"
                     /db_xref="GI:4105186"
                     /translation="MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNC
                     NGWRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEA
                     RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWK
                     GKLRPVPSKLTDLPFFDLMSIPGKLRAGFGAIGLRPSPPGHEESVEQFVRRNLGGEVF
                     ERLIEPFCSGVYAGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRD
                     PRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPE
                     GVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDER
                     LVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKT
                     ESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNG
                     LEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK"
BASE COUNT      509 a    383 c    461 g    521 t
ORIGIN


alignment_scores:
          Quality: 2832.00             Length:     548
            Ratio:  5.177                Gaps:       0
 Percent Similarity:  99.818   Percent Identity: 99.635

alignment_block:
 US-09-508-418-2 x AF044128   ..

 Align seg 1/1  to: AF044128  from: 1  to: 1874

      1 MetThrThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSe 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     26 ATGACAACAACTCCCATCGCCAATCATCCTAATATTTTCACTCACCAGTC 75

     17 rSerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     76 GTCGTCATCGCCATTGGCATTCTTAAACCGTACGAGTTTCATCCCTTTCT 125

     34 erSerIleSerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArg 50
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 126 CTTCAATCTCCAAGCGCAATAGTGTCAATTGCAATGGCTGGAGAACACGA 175

  51 CysSerValAlaLysAspTyrThrValProSerSerAlaValAspGlyGl 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 176 TGCTCCGTTGCCAAAGATTACACAGTTCCTTCCTCAGCGGTCGACGGCGG 225

  67 yProAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyL 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 226 ACCCGCCGCGGAGCTGGACTGTGTTATAGTTGGAGCAGGAATTAGTGGCC 275

  84 euCysIleAlaGlnValMetSerAlaAsnTyrProAsnLeuMetValThr 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 276 TCTGCATTGCGCAGGTGATGTCCGCTAATTACCCCAATTTGATGGTAACC 325

 101 GluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgAspGl 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 326 GAGGCGAGAGATCGTGCCGGTGGCAACATAACGACTGTGGAAAGAGACGG 375

 117 yTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetL 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 376 CTATTTGTGGGAAGAAGGTCCCAACAGTTTCCAGCCGTCCGATCCTATGT 425

 134 euThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyAsp 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 426 TGACTATGGCAGTAGATTGTGGATTGAAGGATGATTTGGTGTTGGGAGAT 475

 151 ProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProValPr 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 476 CCTAATGCGCCCCGTTTCGTTTTGTGGAAGGGTAAATTAAGGCCCGTCCC 525

 167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyL 184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 526 CTCAAAACTCACTGATCTTCCCTTTTTGATTTGATGAGCATTCCTGGCA 575

 184 ysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGly 200
     ||||||||||||||||||||   |||||||||||||||||||||||||||
 576 AGTTGAGAGCTGGTTTTGGTGCCATTGGCCTCCGCCCTTCACCTCCAGGT 625

 201 HisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyGluVa 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 626 CATGAGGAATCAGTTGAGCAGTTCGTGCGTCGTAATCTTGGTGGCGAAGT 675

 217 lPheGluArgLeuIleGluProPheCysSerGlyValTyrValGlyAspP 234
     ||||||||||||||||||||||||||||||||||||||||:::|||||||
 676 CTTTGAACGCTTGATAGAACCATTTTGTTCTGGTGTTTATGCTGGTGATC 725

 234 roSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLeuGlu 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 726 CCTCAAAACTGAGTATGAAAGCAGCATTTGGGAAAGTTTGGAAGTTGGAA 775

 251 GluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysGluAr 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 776 GAAACTGGTGGTAGCATTATTGGAGGAACCTTTAAAGCAATAAAGGAGAG 825

 267 gSerSerThrProLysAlaProArgAspProArgLeuProLysProLysG 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 826 ATCCAGTACACCTAAAGCGCCCCGCGATCCGCGTTTACCTAAACCAAAAG 875

 284 lyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAspAla 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 876 GACAGACAGTTGGATCATTCAGGAAGGGTCTCAGAATGCTGCCGGATGCA 925

 301 IleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSerSe 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 926 ATCAGTGCAAGATTGGGAAGCAAATTAAAACTATCATGGAAGCTTTCTAG 975

 317 rIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrProG 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 976 CATTACTAAGTCAGAAAAAGGAGGATATCACTTGACATACGAGACACCAG 1025

 334 luGlyValValSerLeuGlnSerArgSerIleValMetThrValProSer 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1026 AAGGAGTAGTTTCTCTTCAAAGTCGAAGCATTGTCATGACTGTGCCATCC 1075

 351 TyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAl 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1076 TATGTAGCAAGCAACATATTACGTCCTCTTTCGGTTGCCGCAGCAGATGC 1125

 367 aLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrP 384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1126 ACTTTCAAATTTCTACTATCCCCCAGTTGGAGCAGTCACAATTTCATATC 1175

 384 roGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGly 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1176 CTCAAGAAGCTATTCGTGATGAGCGTCTGGTTGATGGTGAACTAAAGGGA 1225

 401 PheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrIl 417
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1226 TTTGGGCAGTTGCATCCACGTACACAGGGAGTGGAAACACTAGGAACGAT 1275

 417 eTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuL 434
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1276 ATATAGTTCATCACTCTTCCCTAACCGTGCCCCAAAAGGTCGGGTGCTAC 1325

 434 euLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThr 450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1326 TCTTGAACTACATTGGAGGAGCAAAAAATCCTGAAATTTTGTCTAAGACG 1375

 451 GluSerGlnLeuValGluValValAspArgAspLeuArgLysMetLeuIl 467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1376 GAGAGCCAACTTGTGGAAGTAGTTGATCGTGACCTCAGAAAAATGCTTAT 1425

 467 eLysProLysAlaGlnAspProLeuValValGlyValArgValTrpProG 484
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1426 AAAACCCAAAGCTCAAGATCCTCTTGTTGTGGGTGTGCGAGTATGGCCAC 1475

 484 lnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAla 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1476 AAGCTATCCCACAGTTTTTGGTTGGTCATCTGGATACGCTAAGTACTGCA 1525

 501 LysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAs 517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1526 AAAGCTGCTATGAATGATAATGGGCTTGAAGGGCTGTTTCTTGGGGGTAA 1575

 517 nTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluV 534
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1576 TTATGTGTCAGGTGTAGCATTGGGGAGGTGTGTTGAAGGTGCTTATGAAG 1625

 534 alAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
     ||||||||||||||||||||||||||||||||||||||||||||
1626 TTGCATCCGAGGTAACAGGATTTCTGTCTCGGTATGCATACAAA 1669
```

```
seq_name: gb_pl1:AB020501

seq_documentation_block:
LOCUS       AB020501     1647 bp    mRNA            PLN       06-FEB-1999
DEFINITION  Nicotiana tabacum mRNA for plastidal protoporphyrinogen oxidase,
            complete cds.
ACCESSION   AB020501
VERSION     AB020501.1  GI:3929921
KEYWORDS    plastidal protoporphyrinogen oxidase.
SOURCE      Nicotiana tabacum (strain:Samsun NN) cDNA to mRNA.
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1647)
  AUTHORS   Che,F. and Watanabe,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-NOV-1998) to the DDBJ/EMBL/GenBank databases.
            Fang-Sik Che, Nara Institute of Science and Technology, Graduate
            School of Biological Sciences; Takayamachou 8916-5, Ikoma, Nara
            630-0101, Japan (E-mail:fsche@bs.aist-nara.ac.jp,
            Tel:81-743-72-5453, Fax:81-743-72-5459)
REFERENCE   2  (sites)
```

```
  AUTHORS   Watanabe,N., Che,F.S., Iwano,M., Takayama,S., Nakano,T., Yoshida,S.
            and Isogai,A.
  TITLE     Molecular characterization of photomixotrophic tobacco cells
            resistant to protoporphyrinogen oxidase-inhibiting herbicides
  JOURNAL   Plant Physiol. 118 (3), 751-758 (1998)
  MEDLINE   99026276
FEATURES             Location/Qualifiers
     source          1. .1647
                     /organism="Nicotiana tabacum"
                     /strain="Samsun NN"
                     /db_xref="taxon:4097"
                     /note="Photomixotrophic cultured cells"
     gene            1. .1647
                     /gene="NtPPOX1"
     CDS             1. .1647
                     /gene="NtPPOX1"
                     /codon_start=1
                     /product="plastidal protoporphyrinogen oxidase"
                     /protein_id="BAA34713.1"
                     /db_xref="GI:3929922"
                     /translation="MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNC
                     NGWRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEA
                     RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWK
                     GKLRPVPSKLTDLAFFDLMSIPGKLRAGFGAIGLRPSPPGHEESVEQFVRRNLGGEVF
                     ERLIEPFCSGVYAGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRD
                     PRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPE
                     GVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTISYPQEAIRDER
                     LVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKT
                     ESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNG
                     LEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK"
BASE COUNT      437 a    349 c    415 g    446 t
ORIGIN


alignment_scores:
           Quality: 2824.00             Length:     548
             Ratio:  5.172                Gaps:       0
 Percent Similarity:  99.635   Percent Identity: 99.453

alignment_block:
 US-09-508-418-2 x AB020501   ..

 Align seg 1/1  to: AB020501  from: 1  to: 1647

       1 MetThrThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 ATGACAACAACTCCCATCGCCAATCATCCTAATATTTTCACTCACCAGTC 50

      17 rSerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 GTCGTCATCGCCATTGGCATTCTTAAACCGTACGAGTTTCATCCCTTTCT 100

      34 erSerIleSerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArg 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 CTTCAATCTCCAAGCGCAATAGTGTCAATTGCAATGGCTGGAGAACACGA 150

      51 CysSerValAlaLysAspTyrThrValProSerSerAlaValAspGlyGl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 TGCTCCGTTGCCAAAGATTACACAGTTCCTTCCTCAGCGGTCGACGGCGG 200

      67 yProAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyL 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 ACCCGCCGCGGAGCTGGACTGTGTTATAGTTGGAGCAGGAATTAGTGGCC 250

      84 euCysIleAlaGlnValMetSerAlaAsnTyrProAsnLeuMetValThr 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 TCTGCATTGCGCAGGTGATGTCCGCTAATTACCCCAATTTGATGGTAACC 300

     101 GluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgAspGl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 GAGGCGAGAGATCGTGCCGGTGGCAACATAACGACTGTGGAAAGAGACGG 350

     117 yTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetL 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 CTATTTGTGGGAAGAAGGTCCCAACAGTTTCCAGCCGTCCGATCCTATGT 400

     134 euThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyAsp 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 TGACTATGGCAGTAGATTGTGGATTGAAGGATGATTTGGTGTTGGGAGAT 450

     151 ProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProValPr 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     451 CCTAATGCGCCCCGTTTCGTTTTGTGGAAGGGTAAATTAAGGCCCGTCCC 500

     167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyL 184
         |||||||||||||||||||   ||||||||||||||||||||||||||||
     501 CTCAAAACTCACTGATCTTGCCTTTTTTGATTTGATGAGCATTCCTGGCA 550

     184 ysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGly 200
         ||||||||||||||||||||   |||||||||||||||||||||||||||
     551 AGTTGAGAGCTGGTTTTGGTGCCATTGGCCTCCGCCCTTCACCTCCAGGT 600

     201 HisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyGluVa 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     601 CATGAGGAATCAGTTGAGCAGTTCGTGCGTCGTAATCTTGGTGGCGAAGT 650

     217 lPheGluArgLeuIleGluProPheCysSerGlyValTyrValGlyAspP 234
         ||||||||||||||||||||||||||||||||||||||||:::|||||||
     651 CTTTGAACGCTTGATAGAACCATTTTGTTCTGGTGTTTATGCTGGTGATC 700

     234 roSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLeuGlu 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     701 CCTCAAAACTGAGTATGAAAGCAGCATTTGGGAAAGTTTGGAAGTTGGAA 750

     251 GluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysGluAr 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     751 GAAACTGGTGGTAGCATTATTGGAGGAACCTTTAAAGCAATAAAGGAGAG 800

     267 gSerSerThrProLysAlaProArgAspProArgLeuProLysProLysG 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     801 ATCCAGTACACCTAAAGCGCCCCGCGATCCGCGTTTACCTAAACCAAAAG 850

     284 lyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAspAla 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     851 GACAGACAGTTGGATCATTCAGGAAGGGTCTCAGAATGCTGCCGGATGCA 900

     301 IleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSerSe 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     901 ATCAGTGCAAGATTGGGAAGCAAATTAAAACTATCATGGAAGCTTTCTAG 950

     317 rIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrProG 334
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     951 CATTACTAAGTCAGAAAAAGGAGGATATCACTTGACATACGAGACACCAG 1000

     334 luGlyValValSerLeuGlnSerArgSerIleValMetThrValProSer 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1001 AAGGAGTAGTTTCTCTTCAAAGTCGAAGCATTGTCATGACTGTGCCATCC 1050

     351 TyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAl 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1051 TATGTAGCAAGCAACATATTACGTCCTCTTTCGGTTGCCGCAGCAGATGC 1100

     367 aLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrP 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1101 ACTTTCAAATTTCTACTATCCCCCAGTTGGAGCAGTCACAATTTCATATC 1150

     384 roGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGly 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1151 CTCAAGAAGCTATTCGTGATGAGCGTCTGGTTGATGGTGAACTAAAGGGA 1200

     401 PheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrIl 417
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1201 TTTGGGCAGTTGCATCCACGTACACAGGGAGTGGAAACACTAGGAACGAT 1250

     417 eTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuL 434
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
1251 ATATAGTTCATCACTCTTCCCTAACCGTGCCCCAAAAGGTCGGGTGCTAC 1300

 434 euLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThr 450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1301 TCTTGAACTACATTGGAGGAGCAAAAAATCCTGAAATTTTGTCTAAGACG 1350

 451 GluSerGlnLeuValGluValValAspArgAspLeuArgLysMetLeuIl 467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1351 GAGAGCCAACTTGTGGAAGTAGTTGATCGTGACCTCAGAAAAATGCTTAT 1400

 467 eLysProLysAlaGlnAspProLeuValValGlyValArgValTrpProG 484
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1401 AAAACCCAAAGCTCAAGATCCTCTTGTTGTGGGTGTGCGAGTATGGCCAC 1450

 484 lnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAla 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1451 AAGCTATCCCACAGTTTTTGGTTGGTCATCTGGATACGCTAAGTACTGCA 1500

 501 LysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAs 517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1501 AAAGCTGCTATGAATGATAATGGGCTTGAAGGGCTGTTTCTTGGGGGTAA 1550

 517 nTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluV 534
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1551 TTATGTGTCAGGTGTAGCATTGGGGAGGTGTGTTGAAGGTGCTTATGAAG 1600

 534 alAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
     ||||||||||||||||||||||||||||||||||||||||||||
1601 TTGCATCCGAGGTAACAGGATTTCTGTCTCGGTATGCATACAAA 1644

seq_name: gb_pl3:NTY13465

seq_documentation_block:
LOCUS       NTY13465     1892 bp    mRNA            PLN       08-SEP-1997
DEFINITION  Nicotiana tabacum mRNA for protoporphyrin IX oxidase, ppxI gene.
ACCESSION   Y13465
VERSION     Y13465.1  GI:2370332
KEYWORDS    oxidase; ppxI gene; protoporphyrin IX; tetrapyrrole synthesis
            enzyme.
SOURCE      common tobacco.
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1892)
  AUTHORS   Lermontova,I., Kruse,E., Mock,H.P. and Grimm,B.
  TITLE     Cloning and characterization of a plastidal and a mitochondrial
            isoform of tobacco protoporphyrinogen IX oxidase
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 94 (16), 8895-8900 (1997)
  MEDLINE   97385200
REFERENCE   2  (bases 1 to 1892)
  AUTHORS   Inna,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAY-1997) L. Inna, Institute of Plant Genetics (IPK),
            Dept. of Molecular Cell Biology, Corrensstr.3, 06466 Gatersleben,
            FRG
  REMARK    revised by author
FEATURES             Location/Qualifiers
     source          1. .1892
                     /organism="Nicotiana tabacum"
                     /strain="SR1"
                     /db_xref="taxon:4097"
                     /clone_lib="Lambda ZAP II tobacco library"
     gene            29. .1675
                     /gene="ppxI"
     CDS             29. .1675
                     /gene="ppxI"
                     /EC_number="1.3.3.4"
                     /function="tetrapyrrole biosynthesis"
                     /codon_start=1
                     /product="protoporphyrinogen oxidase"
                     /protein_id="CAA73865.1"
                     /db_xref="GI:2370333"
                     /db_xref="SWISS-PROT:O24163"
                     /translation="MTTTPIANHPNIFTHQSSSSPLAFLNRTSFIPFSSISKRNSVNC
                     NGWRTRCSVAKDYTVPSSAVDGGPAAELDCVIVGAGISGLCIAQVMSANYPNLMVTEA
                     RDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDPNAPRFVLWK
                     GKLRPVPSKLTDLAFFDLMSIPGKLRAGFGAIGLRPSPPGHEESVEQFVRRNLGGEVF
                     ERLIEPFCSGVYAGDPSKLSMKAAFGKVWKLEETGGSIIGGTFKAIKERSSTPKAPRD
                     PRLPKPKGQTVGSFRKGLRMLPDAISARLGSKLKLSWKLSSITKSEKGGYHLTYETPE
                     GVVSLQSRSIVMTVPSYVASNILRPLSVAAADALSNFYYPPVGAVTITYPQEAIRDER
                     LVDGELKGFGQLHPRTQGVETLGTIYSSSLFPNRAPKGRVLLLNYIGGAKNPEILSKT
                     ESQLVEVVDRDLRKMLIKPKAQDPLVVGVRVWPQAIPQFLVGHLDTLSTAKAAMNDNG
                     LEGLFLGGNYVSGVALGRCVEGAYEVASEVTGFLSRYAYK"
BASE COUNT      516 a    383 c    467 g    526 t
ORIGIN

alignment_scores:
         Quality:  2821.00              Length:    548
           Ratio:  5.167                  Gaps:      0
 Percent Similarity:  99.635   Percent Identity:  99.270

alignment_block:
 US-09-508-418-2 x NTY13465   ..

 Align seg 1/1  to: NTY13465  from: 1  to: 1892

   1 MetThrThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSe 17
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  29 ATGACAACAACTCCCATCGCCAATCATCCTAATATTTTCACTCACCAGTC 78

  17 rSerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheS 34
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  79 GTCGTCATCGCCATTGGCATTCTTAAACCGTACGAGTTTCATCCCTTTCT 128

  34 erSerIleSerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArg 50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 129 CTTCAATCTCCAAGCGCAATAGTGTCAATTGCAATGGCTGGAGAACACGA 178

  51 CysSerValAlaLysAspTyrThrValProSerSerAlaValAspGlyGl 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 179 TGCTCCGTTGCCAAAGATTACACAGTTCCTTCCTCAGCGGTCGACGGCGG 228

  67 yProAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyL 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 229 ACCCGCCGCGGAGCTGGACTGTGTTATAGTTGGAGCAGGAATTAGTGGCC 278

  84 euCysIleAlaGlnValMetSerAlaAsnTyrProAsnLeuMetValThr 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 279 TCTGCATTGCGCAGGTGATGTCCGCTAATTACCCCAATTTGATGGTAACC 328

 101 GluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgAspGl 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 329 GAGGCGAGAGATCGTGCCGGTGGCAACATAACGACTGTGGAAAGAGACGG 378

 117 yTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetL 134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 379 CTATTTGTGGGAAGAAGGTCCCAACAGTTTCCAGCCGTCCGATCCTATGT 428

 134 euThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyAsp 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 429 TGACTATGGCAGTAGATTGTGGATTGAAGGATGATTTGGTGTTGGGAGAT 478

 151 ProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProValPr 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 479 CCTAATGCGCCCCGTTTCGTTTTGTGGAAGGGTAAATTAAGGCCCGTCCC 528

 167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyL 184
     |||||||||||||||||||   ||||||||||||||||||||||||||||
 529 CTCAAAACTCACTGATCTTGCCTTTTTTGATTTGATGAGCATTCCTGGCA 578

 184 ysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGly 200
     ||||||||||||||||||||   |||||||||||||||||||||||||||
 579 AGTTGAGAGCTGGTTTTGGTGCCATTGGCCTCCGCCCTTCACCTCCAGGT 628
```

```
 201 HisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyGluVa 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 629 CATGAGGAATCAGTTGAGCAGTTCGTGCGTCGTAATCTTGGTGGCGAAGT 678

 217 lPheGluArgLeuIleGluProPheCysSerGlyValTyrValGlyAspP 234
     ||||||||||||||||||||||||||||||||||||||||:::|||||||
 679 CTTTGAACGCTTGATAGAACCATTTTGTTCTGGTGTTTATGCTGGTGATC 728

 234 roSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLeuGlu 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 729 CCTCAAAACTGAGTATGAAAGCAGCATTTGGGAAAGTTTGGAAGTTGGAA 778

 251 GluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysGluAr 267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 779 GAAACTGGTGGTAGCATTATTGGAGGAACCTTTAAAGCAATAAAGGAGAG 828

 267 gSerSerThrProLysAlaProArgAspProArgLeuProLysProLysG 284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 829 ATCCAGTACACCTAAAGCGCCCCGCGATCCGCGTTTACCTAAACCAAAAG 878

 284 lyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAspAla 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 879 GACAGACAGTTGGATCATTCAGGAAGGGTCTCAGAATGCTGCCGGATGCA 928

 301 IleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSerSe 317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 929 ATCAGTGCAAGATTGGGAAGCAAATTAAAACTATCATGGAAGCTTTCTAG 978

 317 rIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrProG 334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 979 CATTACTAAGTCAGAAAAAGGAGGATATCACTTGACATACGAGACACCAG 1028

 334 luGlyValValSerLeuGlnSerArgSerIleValMetThrValProSer 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1029 AAGGAGTAGTTTCTCTTCAAAGTCGAAGCATTGTCATGACTGTGCCATCC 1078

 351 TyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAl 367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1079 TATGTAGCAAGCAACATATTACGTCCTCTTTCGGTTGCCGCAGCAGATGC 1128

 367 aLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrP 384
     |||||||||||||||||||||||||||||||||||||||||||:::||||
1129 ACTTTCAAATTTCTACTATCCCCCAGTTGGAGCAGTCACAATTACATATC 1178

 384 roGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGly 400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1179 CTCAAGAAGCTATTCGTGATGAGCGTCTGGTTGATGGTGAACTAAAGGGA 1228

 401 PheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrIl 417
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1229 TTTGGGCAGTTGCATCCACGTACACAGGGAGTGGAAACACTAGGAACGAT 1278

 417 eTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuL 434
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1279 ATATAGTTCATCACTCTTCCCTAACCGTGCCCCAAAAGGTCGGGTGCTAC 1328

 434 euLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThr 450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1329 TCTTGAACTACATTGGAGGAGCAAAAAATCCTGAAATTTTGTCTAAGACG 1378

 451 GluSerGlnLeuValGluValValAspArgAspLeuArgLysMetLeuIl 467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1379 GAGAGCCAACTTGTGGAAGTAGTTGATCGTGACCTCAGAAAAATGCTTAT 1428

 467 eLysProLysAlaGlnAspProLeuValValGlyValArgValTrpProG 484
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1429 AAAACCCAAAGCTCAAGATCCTCTTGTTGTGGGTGTGCGAGTATGGCCAC 1478

 484 lnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAla 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1479 AAGCTATCCCACAGTTTTTGGTTGGTCATCTGGATACGCTAAGTACTGCA 1528

 501 LysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAs 517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1529 AAAGCTGCTATGAATGATAATGGGCTTGAAGGGCTGTTTCTTGGGGGTAA 1578

 517 nTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluV 534
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1579 TTATGTGTCAGGTGTAGCATTGGGGAGGTGTGTTGAAGGTGCTTATGAAG 1628

 534 alAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
     |||||||||||||||||||||||||||||||||||||||||||||
1629 TTGCATCCGAGGTAACAGGATTTCTGTCTCGGTATGCATACAAA 1672
```

```
seq_name: gb_pl4:STA225107

seq_documentation_block:
LOCUS       STA225107    1976 bp    mRNA            PLN       09-FEB-1999
DEFINITION  Solanum tuberosum (cultivar Bintje) chloroplastic
            protoporphyrinogen IX oxidase.
ACCESSION   AJ225107
VERSION     AJ225107.1  GI:3093409
KEYWORDS    protoporphyrinogen oxidase.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 1976)
  AUTHORS   Johnston,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-1998) Johnston D.J., Biologie moleculaire, RAC,
            Changins, 1260, Nyon, SWITZERLAND
REFERENCE   2  (bases 1 to 1976)
  AUTHORS   Johnston,D.J., Droz,E., Rochaix,J.D. and Malno,P.
  TITLE     Cloning and Characterization of Potato cDNAs Involved in
            Tetrapyrrole Biosynthesis:ferrochelatase (Accession No. AJ005802),
            chloroplastic protoporphyrinogen IX oxidase (Accession No.
            AJ225107), andmitochondrial protoporphyrinogen IX oxidase
            (Accession No. AJ225108)
  JOURNAL   Plant Physiol. 118, 330-330 (1998)
FEATURES             Location/Qualifiers
     source          1. .1976
                     /organism="Solanum tuberosum"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
     CDS             1. .1674
                     /EC_number="1.3.3.4"
                     /function="tetrapyrrole biosynthesis"
                     /codon_start=1
                     /product="protoporphyrinogen oxidase"
                     /protein_id="CAA12400.1"
                     /db_xref="GI:3093410"
                     /db_xref="SPTREMBL:O64384"
                     /translation="MTTTAVANHPSIFTHRSPLPSPSSSSSSPSFLFLNRTNFIPYFS
                     TSKRNSVNCNGWRTRCSVAKDYTVPPSEVDGNQFPELDCVVVGAGISGLCIAKVISAN
                     YPNLMVTEARDRAGGNITTVERDGYLWEEGPNSFQPSDPMLTMAVDCGLKDDLVLGDP
                     DAPRFVLWKDKLRPVPGKLTDLPFFDLMSIPGKLRAGFGAIGLRPSPPGYEESVEQFV
                     RRNLGAEVFERLIEPFCSGVYAGDPSKLIMKAAFGKVWKLEQTGGSIIGGTFKAIKER
                     SSNPKPPRDPRLPTPKGQTVGSFRKGLRMLPDAICERLGSKVKLSWKLSSITKSEKGG
                     YLLTYETPEGVVSLRSRSIVMTVPSYVASNILRPLSVAAADALSSFYYPPVAAVTISY
                     PQEAIRDERLVDGELKGFGQLHPRSQGVETLGTIYSSSLFPNRAPNGRVLLLNYIGGA
                     TNTEIVSKTESQLVEAVDRDLRKMLIKPKAQDPFVTGVRVWPQAIPQFLVGHLDTLGT
                     AKTALSDNGLDGLFLGGNYVSGVALGRCVEGAYEIASEVTGFLSQYAYK"
     3'UTR           1675. .1976
BASE COUNT      518 a    394 c    485 g    579 t
ORIGIN


alignment_scores:
           Quality: 2565.50            Length:    557
             Ratio:  4.850               Gaps:      2
 Percent Similarity:  94.973   Percent Identity: 88.869

alignment_block:
 US-09-508-418-2 x STA225107   ..
```

```
Align seg 1/1  to: STA225107  from: 1  to: 1976

    1 MetThrThrThrProIleAlaAsnHisProAsnIlePheThrHisGln.. 16
      ||||||||||||   :::||||||||||||:::||||||||||||:::
    1 ATGACAACAACGGCCGTCGCCAACCATCCTAGCATTTTCACTCACCGGTC 50

   17 ......................SerSerSerSerPro...LeuAlaPheL 25
                            |||||||||||||||   :::   ||||
   51 GCCGCTGCCGTCGCCGTCGTCCTCCTCCTCATCGCCGTCATTTTTATTTT 100

   25 euAsnArgThrSerPheIleProPheSerSerIleSerLysArgAsnSer 41
      |||||||||||:::|||||||||:::   |||   |||||||||||||||
  101 TAAACCGTACGAATTTCATTCCTTACTTTTCCACCTCCAAGCGCAATAGT 150

   42 ValAsnCysAsnGlyTrpArgThrArgCysSerValAlaLysAspTyrTh 58
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  151 GTCAATTGCAATGGCTGGAGAACACGATGTTCCGTTGCCAAGGATTATAC 200

   58 rValProSerSerAlaValAspGlyGlyProAlaAlaGluLeuAspCysV 75
      |||||||   |||   |||||||||:::         |||||||||||||
  201 AGTTCCTCCCTCGGAAGTCGACGGTAATCAGTTCCCGGAGCTGGATTGTG 250

   75 alIleValGlyAlaGlyIleSerGlyLeuCysIleAlaGlnValMetSer 91
      ||:::|||||||||||||||||||||||||||||||||:::|||:::|||
  251 TGGTAGTTGGAGCAGGAATTAGTGGACTCTGCATTGCTAAGGTGATTTCG 300

   92 AlaAsnTyrProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGl 108
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  301 GCTAATTATCCCAATTTGATGGTGACGGAGGCGAGGGATCGTGCCGGTGG 350

  108 yAsnIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProA 125
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  351 AAACATAACGACGGTGGAAAGAGATGGATACTTATGGGAAGAAGGTCCTA 400

  125 snSerPheGlnProSerAspProMetLeuThrMetAlaValAspCysGly 141
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  401 ACAGTTTCCAGCCTTCGGATCCTATGTTGACAATGGCTGTAGATTGTGGA 450

  142 LeuLysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLe 158
      ||||||||||||||||||||||||||||||:::|||||||||||||||||
  451 TTGAAGGATGATTTGGTGTTGGGAGATCCTGATGCGCCTCGCTTTGTCTT 500

  158 uTrpLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProP 175
      |||||||   ||||||||||||||||||:::|||||||||||||||||||
  501 GTGGAAGGATAAACTAAGGCCTGTTCCCGGCAAGCTCACTGATCTTCCCT 550

  175 hePheAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyPro 191
      |||||||||||||||||||||||||||||||||||||||||||||||
  551 TCTTTGATTTGATGAGTATCCCTGGCAAGCTCAGAGCTGGTTTTGGTGCC 600

  192 IleGlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPh 208
      |||||||||||||||||||||||||||:::||||||||||||||||||||
  601 ATTGGCCTTCGCCCTTCACCTCCAGGTTATGAGGAATCAGTTGAGCAGTT 650

  208 eValArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProP 225
      |||||||||||||||||||:::||||||||||||||||||||||||||||
  651 CGTGCGTCGTAATCTTGGTGCAGAAGTCTTTGAACGTTTGATTGAACCAT 700

  225 heCysSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAla 241
      |||||||||||||||||:::||||||||||||||||||   |||||||||
  701 TTTGTTCTGGTGTTTACGCCGGTGACCCCTCAAAATTGATTATGAAAGCA 750

  242 AlaPheGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGl 258
      |||||||||||||||||||||||||||:::||||||||||||||||||||
  751 GCATTTGGGAAAGTGTGGAAGCTAGAACAAACTGGTGGTAGCATTATTGG 800

  258 yGlyThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProA 275
      ||||||||||||||||||||||||||||||||||:::||||||   ||||
  801 GGGAACCTTTAAAGCAATTAAGGAGAGATCCAGTAACCCTAAACCGCCTC 850

  275 rgAspProArgLeuProLysProLysGlyGlnThrValGlySerPheArg 291
      |||||||||||||||||   ||||||||||||||||||||||||||||||
  851 GTGATCCGCGTTTACCAACACCAAAAGGACAAACTGTTGGATCATTTAGG 900

  292 LysGlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLy 308
      ||||||||||||||||||||||||||||||      ||||||||||||||
  901 AAGGGTCTGAGAATGCTGCCGGATGCAATTTGTGAAAGACTGGGAAGCAA 950

  308 sLeuLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyG 325
      |:::||||||||||||||||||||||||||||||||||||||||||||||
  951 AGTAAAACTATCATGGAAGCTTTCTAGCATTACAAAGTCAGAAAAAGGAG 1000

  325 lyTyrHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSer 341
      |||||   ||||||||||||||||||||||||||||||||||||:::|||
 1001 GATATCTCTTGACATACGAGACACCAGAAGGAGTAGTTTCTCTGCGAAGT 1050

  342 ArgSerIleValMetThrValProSerTyrValAlaSerAsnIleLeuAr 358
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1051 CGAAGCATTGTCATGACTGTTCCATCCTATGTAGCAAGCAACATATTACG 1100

  358 gProLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProP 375
      ||||||||||||||||||||||||||||||||||:::|||||||||||||
 1101 CCCTCTTTCGGTCGCTGCAGCAGATGCACTTTCAAGTTTCTACTATCCCC 1150

  375 roValGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGlu 391
      |||||:::||||||||||||||||||||||||||||||||||||||||||
 1151 CAGTAGCAGCAGTGACAATTTCATATCCTCAAGAGGCTATTCGTGATGAG 1200

  392 ArgLeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgTh 408
      ||||||||||||||||||||||||||||||||||||||||||||||||::
 1201 CGTCTGGTTGATGGTGAACTAAAGGGATTTGGGCAGTTGCATCCACGTTC 1250

  408 rGlnGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProA 425
      :|||||||||||||||||||||||||||||||||||||||||||||||||
 1251 ACAGGGAGTGGAAACACTAGGAACAATATATAGTTCATCACTCTTTCCTA 1300

  425 snArgAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAla 441
      |||||||||||:::||||||||||||||||||||||||||||||||||||
 1301 ACCGTGCTCCAAATGGCCGGGTGCTACTCTTGAACTACATTGGAGGAGCA 1350

  442 LysAsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValVa 458
         |||   ||||||:::|||||||||||||||||||||||||||:::||
 1351 ACAAATACTGAAATTGTGTCTAAGACGGAGAGCCAACTTGTGGAAGCAGT 1400

  458 lAspArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProL 475
      |||||||||||||||||||||||||||||||||||||||||||||||||:
 1401 TGACCGTGACCTCAGAAAAATGCTTATAAAACCCAAAGCACAAGATCCCT 1450

  475 euValValGlyValArgValTrpProGlnAlaIleProGlnPheLeuVal 491
      ::|||:::||||||||||||||||||||||||||||||||||||||||||
 1451 TTGTTACGGGTGTGCGAGTATGGCCACAAGCTATCCCACAGTTTTTGGTC 1500

  492 GlyHisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGl 508
      ||||||||||||||||||:::|||||||||:::|||::::::||||||||
 1501 GGACATCTGGATACACTAGGTACTGCAAAAACTGCTCTAAGTGATAATGG 1550

  508 yLeuGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuG 525
      ||||:::|||||||||||||||||||||||||||||||||||||||||||
 1551 GCTTGACGGGCTATTCCTTGGGGGTAATTATGTGTCTGGTGTAGCATTGG 1600

  525 lyArgCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPhe 541
      ||||||||||||||||||||||||||:::|||||||||||||||||||||
 1601 GAAGGTGTGTTGAAGGTGCTTATGAAATAGCATCTGAGGTAACTGGATTT 1650

  542 LeuSerArgTyrAlaTyrLys 548
      ||||||:::||||||||||||
 1651 CTGTCTCAGTATGCATACAAA 1671

seq_name: gb_pat2:AX084746

seq_documentation_block:
LOCUS       AX084746     1826 bp    DNA             PAT       09-MAR-2001
DEFINITION  Sequence 15 from Patent WO0112825.
ACCESSION   AX084746
VERSION     AX084746.1  GI:13274999
```

```
KEYWORDS    .
SOURCE      upland cotton.
  ORGANISM  Gossypium hirsutum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Malvales; Malvaceae; Gossypium.
REFERENCE   1  (bases 1 to 1826)
  AUTHORS   Johnson,M.A., Volrath,S.L., Heifetz,P.B. and Law,M.D.
  TITLE     Herbicide-tolerant protoporphyrinogen oxidase
  JOURNAL   Patent: WO 0112825-A 15 22-FEB-2001;
            Syngenta Participations AG (CH)
FEATURES             Location/Qualifiers
     source          1..1826
                     /organism="Gossypium hirsutum"
                     /db_xref="taxon:3635"
     CDS             31..1650
                     /note="cotton protox-1 coding sequence"
                     /codon_start=1
                     /protein_id="CAC34014.1"
                     /db_xref="GI:13275000"
                     /translation="MTALIDLSLLRSSPSVSPFSIPHHQHPPRFRKPFKLRCSLAEGP
                     TISSSKIDGGESSIADCVIVGGGISGLCIAQALATKHRDVASNVIVTEARDRVGGNIT
                     TVERDGYLWEEGPNSFQPSDPILTMAVDSGLKDDLVLGDPNAPRFVLWEGKLRPVPSK
                     PTDLPFFDLMSIAGKLRAGFGAIGIRPPPPGYEESVEEFVRRNLGAEVFERFIEPFCS
                     GVYAGDPSKLSMKAAFGRVWKLEEIGGSIIGGTFKTIQERNKTPKPPRDPRLPKPKGQ
                     TVGSFRKGLTMLPEAIANSLGSNVKLSWKLSSITKLGNGGYNLTFETPEGMVSLQSRS
                     VVMTIPSHVASNLLHPLSAAAADALSQFYYPPVASVTVSYPKEAIRKECLIDGELKGF
                     GQLHPRSQGIETLGTIYSSSLFPNRAPSGRVLLLNYIGGATNTGILSKTEGELVEAVD
                     RDLRKMLINPNAKDPLVLGVRVWPKAIPQFLVGHLDLLDSAKMALRDSGFHGLFLGGN
                     YVSGVALGRCVEGAYEVAAEVKEFLSQYAYK"
BASE COUNT      471 a    399 c    439 g    517 t
ORIGIN


alignment_scores:
           Quality: 2113.00            Length:    506
             Ratio:  4.477               Gaps:      1
 Percent Similarity:  93.281   Percent Identity: 79.249

alignment_block:
 US-09-508-418-2 x AX084746   ..

 Align seg 1/1  to: AX084746  from: 1  to: 1826

  47 TrpArgThrArgCysSerValAlaLysAspTyrThrValProSerSerAl 63
     ::::::   |||||||||:::|||:::      |||:::   ||||||
 130 TTCAAGCTCCGATGCTCCCTCGCCGAGGGTCCCACGATTTCCTCATCTAA 179

  63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
      :::|||||||||   ::::::      ||||||||||||||||||:::|
 180 AATCGACGGGGGAGAATCATCCATCGCGGATTGCGTCATCGTTGGAGGTG 229

  80 lyIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr...... 94
     ||||||||||||||||||||||||||::::::::::::::::::
 230 GTATCAGTGGACTTTGCATTGCTCAAGCTCTCGCCACCAAGCACCGTGAC 279

  95 ......ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAs 109
              |||::::::|||||||||||||||||||||:::||||||||
 280 GTCGCTTCCAATGTGATTGTGACGGAGGCCAGAGACCGTGTTGGTGGCAA 329

 109 nIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnS 126
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 330 CATCACTACCGTTGAGAGAGATGGATATCTGTGGGAAGAAGGCCCCAACA 379

 126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
     ||||||||||||||||||||:::||||||||||||||||||   ||||||
 380 GTTTTCAGCCCTCCGATCCTATTCTAACCATGGCCGTGGATAGTGGATTG 429

 143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 430 AAGGACGATTTGGTTTTAGGTGACCCTAATGCACCGCGATTTGTACTATG 479

 159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
     |:::|||||||||||||||||||||||||||   ||||||||||||||||
 480 GGAGGGAAAACTAAGGCCTGTGCCCTCCAAGCCAACCGACTTGCCGTTTT 529

 176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
     |||||||||||||||||   ||||||||||||||||||||||||   |||
 530 TTGATTTGATGAGCATTGCTGGAAAACTTAGGGCTGGGTTCGGGGCTATT 579

 193 GlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVa 209
     |||:::||||||   |||||||||:::|||||||||||||||:::|||||
 580 GGCATTCGGCCTCCCCCTCCGGGTTATGAAGAATCGGTGGAGGAGTTTGT 629

 209 lArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC 226
     ||||||||||||||||:::|||||||||||||||:::|||||||||||||
 630 GCGCCGTAATCTTGGTGCTGAGGTTTTTGAACGCTTTATTGAACCATTTT 679

 226 ysSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
     ||||||||||||||:::|||||||||||||||||||||||||||||||||
 680 GTTCAGGTGTTTATGCAGGGGATCCTTCAAAATTAAGCATGAAAGCAGCA 729

 243 PheGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGl 259
     ||||||:::||||||||||||||||||   ||||||||||||||||||||
 730 TTTGGAAGAGTATGGAAGCTAGAAGAGATTGGTGGCAGCATCATTGGTGG 779

 259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
     ||||||||||:::|||:::||||||::::::|||||||||   |||||||
 780 CACTTTCAAGACAATCCAGGAGAGAAATAAGACACCTAAGCCACCCAGAG 829

 276 spProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 292
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 830 ACCCGCGTCTGCCAAAACCGAAGGGCCAAACAGTTGGATCTTTTAGGAAG 879

 293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
     ||||||   |||||||||:::||||||:::      |||||||||:::::
 880 GGACTTACCATGCTGCCTGAGGCAATTGCTAACAGTTTGGGTAGCAATGT 929

 309 uLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyT 326
     :|||||||||||||||||||||||||||||||||      :::|||||||
 930 AAAATTATCTTGGAAGCTTTCCAGTATTACCAAATTGGGCAATGGAGGGT 979

 326 yrHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
     ||:::||||||:::|||||||||||||||:::||||||||||||||||||
 980 ATAACTTGACATTTGAAACACCTGAAGGAATGGTATCTCTTCAGAGTAGA 1029

 343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
     |||:::|||||||||:::||||||:::||||||||||||:::|||:::||
1030 AGTGTTGTAATGACCATTCCATCCCATGTTGCCAGTAACTTGTTGCATCC 1079

 359 oLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProV 376
     |||||||:::|||||||||||||||||||||:::||||||||||||||||
1080 TCTCTCGGCTGCTGCTGCAGATGCATTATCCCAATTTTATTATCCTCCAG 1129

 376 alGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArg 392
     ||::::::||||||:::|||||||||:::||||||||||||   |||
1130 TTGCATCAGTCACAGTCTCCTATCCAAAAGAAGCCATTCGAAAAGAATGT 1179

 393 LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGl 409
     |||:::|||||||||||||||||||||||||||||||||||||||:::||
1180 TTGATTGATGGTGAACTTAAGGGGTTTGGCCAGTTGCACCCACGCAGCCA 1229

 409 nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnA 426
     ||||:::|||||||||||||||||||||||||||||||||||||||||||
1230 AGGAATTGAAACTTTAGGGACGATATACAGTTCATCACTTTTCCCCAATC 1279

 426 rgAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
     ||||||||:::||||||||||||||||||||||||||||||||||||
1280 GAGCTCCATCTGGCAGGGTGTTGCTCTTGAACTACATAGGAGGAGCTACC 1329

 443 AsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAs 459
     |||      ||||||||||||||||||::::::|||||||||:::|||||
1330 AACACTGGAATTTTGTCCAAGACTGAAGGGGAACTTGTAGAAGCAGTTGA 1379

 459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
     |||||||||||||||||||||||||:::|||:::|||:::||||||||||
```

```
 1380 TCGTGATTTGAGAAAAATGCTTATAAATCCTAATGCAAAGGATCCTCTTG 1429

  476 alValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGly 492
      ||:::||||||||||||||||||:::||||||||||||||||||||||||
 1430 TTTTGGGTGTAAGAGTATGGCCAAAAGCCATTCCACAGTTCTTGGTTGGT 1479

  493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe 509
      |||||||||   |||::::::||||||   |||::::::|||:::|||::
 1480 CATTTGGATCTCCTTGATAGTGCAAAAATGGCTCTCAGGGATTCTGGGTT 1529

  509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
      ::::||||||||||||||||||||||||||||||||||||||||||||||
 1530 TCATGGACTGTTTCTTGGGGGCAACTATGTATCTGGTGTGGCATTAGGAC 1579

  526 rgCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 542
      |||||||||||||||||||||||||||||:::||||||      ||||||
 1580 GGTGTGTGGAAGGTGCTTACGAGGTTGCAGCTGAAGTGAAGGAATTCCTG 1629

  543 SerArgTyrAlaTyrLys 548
      |||:::||||||||||||
 1630 TCACAATATGCATACAAA 1647

seq_name: gb_pat1:A78957

seq_documentation_block:
LOCUS       A78957       1704 bp    DNA             PAT       20-OCT-1999
DEFINITION  Sequence 11 from Patent WO9829554.
ACCESSION   A78957
VERSION     A78957.1  GI:6092085
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1704)
  AUTHORS   Ishige,F. and Sato,R.
  TITLE     METHODS OF CONFERRING PPO-INHIBITING HERBICIDE RESISTANCE TO PLANTS
            BY GENE MANIPULATION
  JOURNAL   Patent: WO 9829554-A 11 09-JUL-1998;
            ISHIGE FUMIHARU (JP); SATO RYO (JP)
FEATURES             Location/Qualifiers
     source          1..1704
                     /organism="Arabidopsis thaliana"
                     /strain="ECOTYPE COLUMBIA"
                     /db_xref="taxon:3702"
     CDS             16..1629
                     /codon_start=1
                     /product="PROTOPORPHYRINOGEN OXIDASE"
                     /protein_id="CAB59132.1"
                     /db_xref="GI:6092086"
                     /translation="MELSLLRPTTQSLLPSFSKPNLRLNVYKPLRLRCSVAGGPTVGS
                     SKIEGGGGTTITTDCVIVGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITR
                     EENGFLWEEGPNSFQPSDPMLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLT
                     DLPFFDLMSIGGKIRAGFGALGIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGV
                     YAGDPSKLSMKAAFGKVWKLEQNGGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTV
                     GSFRKGLRMLPEAISARLGSKVKLSWKLSGITKLESGGYNLTYETPDGLVSVQSKSVV
                     MTVPSHVASGLLRPLSESAANALSKLYYPPVAAVSISYPKEAIRTECLIDGELKGFGQ
                     LHPRTQGVETLGTIYSSSLFPNRAPPGRILLLNYIGGSTNTGILSKSEGELVEAVDRD
                     LRKMLIKPNSTDPLKLGVRVWPQAIPQFLVGHFDILDTAKSSLTSSGYEGLFLGGNYV
                     AGVALGRCVEGAYETAIEVNNFMSRYAYK"
BASE COUNT      453 a    348 c    431 g    472 t
ORIGIN


alignment_scores:
            Quality:  2077.00              Length:    539
              Ratio:  4.230                  Gaps:      4
 Percent Similarity:  91.095    Percent Identity: 75.139

alignment_block:
 US-09-508-418-2 x A78957   ..
```

```
Align seg 1/1  to: A78957  from: 1  to: 1704

   19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
      |||      |||::::::|||:::   |||:::         :::   ||
   13 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 62

   35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
      |:::||||||   |||   ::::::   |||   :::         |||
   63 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 109

   49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
        |||||||||||||||         ||||||   ||||||   ::::::
  110 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 159

   66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
      ||||||   ::::::   :::      ||||||||||||||||||:::||
  160 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 209

   80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
      |||||||||||||||||||||||||::::::::::::::::::
  210 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 259

   95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
           |||||||||:::||||||||||||:::||||||:::|||||||||
  260 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 309

  110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
      |||   |||   |||::::::|||:::|||||||||||||||||||||||
  310 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 359

  126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
      |||||||||||||||||||||||||||||||:::||||||   |||||||
  360 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 409

  143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
      ||||||||||||||||||||||||||:::|||||||||||||||||||||
  410 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 459

  160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
      :::|||||||||||||||||||||||||||||||||||||||||||||||
  460 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 509

  176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
      ||||||||||||||||   ||||||:::|||||||||||||||   :::|
  510 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 559

  193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
      ||:::||||||||||||||||||:::|||||||||||||||:::||||||
  560 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 609

  210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
      |||||||||||||||   ||||||||||||||||||||||||||||||||
  610 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 659

  226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
      |||||||||||||:::||||||||||||||||||||||||||||||||||
  660 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 709

  243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
      |||||||||||||||||||||||::::::|||||||||||||||||||||
  710 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 759

  260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
      |||||||||||||||:::||||||:::::::::|||||||||   |||||
  760 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 809

  276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
      |||||||||||||||||||:::||||||||||||||||||||||||||||
  810 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 859

  293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
      |||||||||||||||||:::|||||||||||||||||||||||||||:::
  860 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 909
```

```
  310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
      |||||||||||||||||||||:::|||||||||   |||:::||||||||
  910 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 959

  326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
      |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
  960 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1009

  343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
      ||:::||||||||||||||||||:::|||||||||::::::|||||||||
 1010 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1059

  360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
      ||||||   :::||||||:::|||||||||::::::||||||||||||||
 1060 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1109

  376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
      |:::||||||:::||||||||||||:::||||||||||||   |||   |
 1110 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1159

  393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
      ||:::|||||||||||||||||||||||||||||||||||||||||||||
 1160 TGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAA 1209

  410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1210 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1259

  426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
      |||||||   ||||||:::||||||||||||||||||||||||:::   |
 1260 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1309

  443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
      ||      ||||||||||||:::|||::::::|||||||||:::||||||
 1310 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1359

  460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
      ||||||||||||||||||||||||||||||::::::   |||||||||
 1360 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1409

  476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
       :::||||||||||||||||||||||||||||||||||||||||||||||
 1410 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1459

  493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
      ||:::|||   |||:::|||||||||::::::::::::::::::|||
 1460 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1509

  510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
      ||||||||||||||||||||||||||||||:::|||||||||||||||||
 1510 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1559

  526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
      ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
 1560 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1609

  543 erArgTyrAlaTyrLys 548
      |||||||||||||||||
 1610 CACGGTACGCTTACAAG 1626

seq_name: gb_pat2:AX084732

seq_documentation_block:
LOCUS       AX084732     1719 bp    DNA             PAT       09-MAR-2001
DEFINITION  Sequence 1 from Patent WO0112825.
ACCESSION   AX084732
VERSION     AX084732.1  GI:13274985
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1719)
  AUTHORS   Johnson,M.A., Volrath,S.L., Heifetz,P.B. and Law,M.D.
  TITLE     Herbicide-tolerant protoporphyrinogen oxidase
  JOURNAL   Patent: WO 0112825-A 1 22-FEB-2001;
            Syngenta Participations AG (CH)
FEATURES             Location/Qualifiers
     source          1. .1719
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
     CDS             31. .1644
                     /note="Arabidopsis protox-1"
                     /codon_start=1
                     /protein_id="CAC34008.1"
                     /db_xref="GI:13274986"
                     /translation="MELSLLRPTTQSLLPSFSKPNLRLNVYKPLRLRCSVAGGPTVGS
                     SKIEGGGGTTITTDCVIVGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITR
                     EENGFLWEEGPNSFQPSDPMLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLT
                     DLPFFDLMSIGGKIRAGFGALGIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGV
                     YAGDPSKLSMKAAFGKVWKLEQNGGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTV
                     GSFRKGLRMLPEAISARLGSKVKLSWKLSGITKLESGGYNLTYETPDGLVSVQSKSVV
                     MTVPSHVASGLLRPLSESAANALSKLYYPPVAAVSISYPKEAIRTECLIDGELKGFGQ
                     LHPRTQGVETLGTIYSSSLFPNRAPPGRILLLNYIGGSTNTGILSKSEGELVEAVDRD
                     LRKMLIKPNSTDPLKLGVRVWPQAIPQFLVGHFDILDTAKSSLTSSGYEGLFLGGNYV
                     AGVALGRCVEGAYETAIEVNNFMSRYAYK"
BASE COUNT      460 a    351 c    433 g    475 t
ORIGIN


alignment_scores:
           Quality: 2077.00            Length:    539
             Ratio:  4.230               Gaps:      4
 Percent Similarity:  91.095   Percent Identity: 75.139

alignment_block:
 US-09-508-418-2 x AX084732    ..

 Align seg 1/1  to: AX084732  from: 1  to: 1719

         19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
            |||      |||::::::|||:::   |||:::         :::   ||
         28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

         35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
            |:::||||||   |||   ::::::   |||   :::         |||
         78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

         49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
              |||||||||||||||         ||||||   ||||||   ::::::
        125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

         66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
            ||||||   ::::::   :::      ||||||||||||||||||:::||
        175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

         80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
            |||||||||||||||||||||||||::::::::::::::::::
        225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274

         95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
                 |||||||||:::||||||||||||:::||||||:::|||||||||
        275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324

        110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
            |||   |||   |||::::::|||:::|||||||||||||||||||||||
        325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

        126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
            |||||||||||||||||||||||||||||||:::||||||   |||||||
        375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

        143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
            ||||||||||||||||||||||||||:::|||||||||||||||||||||
```

```
 425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

 160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
     :::|||||||||||||||||||||||||||||||||||||||||||||||
 475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524

 176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
     ||||||||||||||||   ||||||:::|||||||||||||||   :::|
 525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

 193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
     ||:::||||||||||||||||||:::|||||||||||||||:::||||||
 575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

 210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
     |||||||||||||||   ||||||||||||||||||||||||||||||||
 625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674

 226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
     |||||||||||||:::||||||||||||||||||||||||||||||||||
 675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

 243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
     |||||||||||||||||||||||::::::|||||||||||||||||||||
 725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

 260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
     |||||||||||||||:::||||||:::::::::|||||||||   |||||
 775 ACTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

 276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
     |||||||||||||||||||:::||||||||||||||||||||||||||||
 825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 874

 293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
     |||||||||||||||||:::|||||||||||||||||||||||||||:::
 875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

 310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
     |||||||||||||||||||||:::|||||||||   |||:::||||||||
 925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

 326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
     |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
 975 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1024

 343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
     ||:::||||||||||||||||||:::|||||||||::::::|||||||||
1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
     ||||||   :::||||||:::|||||||||::::::||||||||||||||
1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

 376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
     |:::||||||:::||||||||||||:::||||||||||||   |||   |
1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

 393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
     ||:::|||||||||||||||||||||||||||||||||||||||||||||
1175 TGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAA 1224

 410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274

 426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
     |||||||   ||||||:::||||||||||||||||||||||||:::   |
1275 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

 443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
     ||      ||||||||||||:::|||::::::|||||||||:::||||||
1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

 460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
     ||||||||||||||||||||||||||||||::::::   |||||||||
1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

 476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
      :::||||||||||||||||||||||||||||||||||||||||||||||
1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

 493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
     ||:::|||   |||:::|||||||||::::::::::::::::::|||
1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

 510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
     ||||||||||||||||||||||||||||||:::|||||||||||||||||
1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

 526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
     ||||||||||||||||||||||:::|||   |||||::::::|||::::|
1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

 543 erArgTyrAlaTyrLys 548
     |||||||||||||||||
1625 CACGGTACGCTTACAAG 1641
```

```
seq_name: gb_pl2:ATHPPOX

seq_documentation_block:
LOCUS       ATHPPOX      1614 bp    mRNA            PLN       06-FEB-1999
DEFINITION  Arabidopsis thaliana mRNA for protoporphyrinogen oxidase, complete
            cds.
ACCESSION   D83139
VERSION     D83139.1  GI:1183006
KEYWORDS    protoporphyrinogen oxidase; PPOX.
SOURCE      Arabidopsis thaliana cDNA to mRNA.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1614)
  AUTHORS   Narita,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JAN-1996) to the DDBJ/EMBL/GenBank databases.
            Shin-ichiro Narita, Kyoto University, Department of Biophysics,
            Faculty of Science; Kitashirakawa-Oiwake-cho, Sakyo-ku, Kyoto,
            Kyoto 606-01, Japan (E-mail:narita@molbio.biophys.kyoto-u.ac.jp,
            Tel:075-753-4201, Fax:075-791-0271)
REFERENCE   2  (bases 1 to 1614)
  AUTHORS   Narita,S., Tanaka,R., Ito,T., Okada,K., Taketani,S. and Inokuchi,H.
  TITLE     Molecular cloning and characterization of a cDNA that encodes
            protoporphyrinogen oxidase of Arabidopsis thaliana
  JOURNAL   Gene 182 (1-2), 169-175 (1996)
  MEDLINE   97136707
FEATURES             Location/Qualifiers
     source          1. .1614
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
     CDS             1. .1614
                     /codon_start=1
                     /product="protoporphyrinogen oxidase"
                     /protein_id="BAA11820.1"
                     /db_xref="GI:1877018"
                     /translation="MELSLLRPTTQSLLPSFSKPNLRLNVYKPLRLRCSVAGGPTVGS
                     SKIEGGGGTTITTDCVIVGGGISGLCIAQALATKHPDAAPNLIVTEAKDRVGGNIITR
                     EENGFLWEEGPNSFQPSDPMLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLT
                     DLPFFDLMSIGGKIRAGFGALGIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGV
                     YAGDPSKLSMKAAFGKVWKLEQNGGSIIGGTFKAIQERKNAPKAERDPRLPKPQGQTV
                     GSFRKGLRMLPEAISARLGSKVKLSWKLSGITKLESGGYNLTYETPDGLVSVQSKSVV
                     MTVPSHVASGLLRPLSESAANALSKLYYPPVAAVSISYPKEAIRTECLIDGELKGFGQ
                     LHPRTQGVETLGTIYSSSLFPNRAPPGRILLLNYIGGSTNTGILSKSEGELVEAVDRD
                     LRKMLIKPNSTDPLKLGVRVWPQAIPQFLVGHFDILDTAKSSLTSSGYEGLFLGGNYV
                     AGVALGRCVEGAYETAIEVNNFMSRYAYK"
BASE COUNT      416 a    334 c    421 g    443 t
```

```
ORIGIN

alignment_scores:
           Quality: 2075.00          Length:    536
             Ratio:  4.235             Gaps:      4
 Percent Similarity:  91.418   Percent Identity: 75.373

alignment_block:
 US-09-508-418-2 x ATHPPOX    ..

 Align seg 1/1  to: ATHPPOX  from: 1  to: 1614

  22 LeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerIleSerLy 38
     |||::::::|||:::   |||:::         :::   |||:::|||||
   7 TTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTCGTTTTCGAA 56

  38 sArgAsnSerValAsnCysAsnGlyTrp.........ArgThrArgCysS 52
     |   |||   ::::::   |||   :::         |||   |||||||
  57 GCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGACTCCGTTGTT 103

  52 erValAlaLysAspTyrThrValProSerSerAlaValAspGlyGlyPro 68
     ||||||||         ||||||   ||||||   ::::::||||||
 104 CAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAAGGCGGAGGA 153

  69 AlaAlaGluLeu......AspCysValIleValGlyAlaGlyIleSerGl 83
     ::::::   :::      ||||||||||||||||||:::|||||||||||
 154 GGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGGTATTAGTGG 203

  83 yLeuCysIleAlaGlnValMetSerAlaAsnTyr............ProA 96
     ||||||||||||||||::::::::::::::::::            ||||
 204 TCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATGCTGCTCCGA 253

  96 snLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThr 112
     |||||:::||||||||||||:::||||||:::||||||||||||   |||
 254 ATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAACATTATCACT 303

 113 ValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnPr 129
        |||::::::|||:::||||||||||||||||||||||||||||||||
 304 CGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAGTTTTCAACC 353

 129 oSerAspProMetLeuThrMetAlaValAspCysGlyLeuLysAspAspL 146
     ||||||||||||||||||||||:::||||||   ||||||||||||||||
 354 GTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGAAGGATGATT 403

 146 euValLeuGlyAspProAsnAlaProArgPheValLeuTrpLysGlyLys 162
     |||||||||||||||||:::|||||||||||||||||||||:::||||||
 404 TGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGGAATGGGAAA 453

 163 LeuArgProValProSerLysLeuThrAspLeuProPhePheAspLeuMe 179
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 454 TTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTTTGATTTGAT 503

 179 tSerIleProGlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgP 196
     |||||||   ||||||:::|||||||||||||||.  :::|||:::||||
 504 GAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTGGCATTCGAC 553

 196 roSerProProGlyHisGluGluSerValGluGlnPheValArgArgAsn 212
     ||||||||||||||:::|||||||||||||||:::|||||||||||||||
 554 CGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTACGGCGTAAC 603

 213 LeuGlyGlyGluValPheGluArgLeuIleGluProPheCysSerGlyVa 229
     ||||||   |||||||||||||||||||||||||||||||||||||||||
 604 CTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTGTTCAGGTGT 653

 229 lTyrValGlyAspProSerLysLeuSerMetLysAlaAlaPheGlyLysV 246
     ||||:::|||||||||||||||||||||||||||||||||||||||||||
 654 TTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGTTTGGGAAGG 703

 246 alTrpLysLeuGluGluThrGlyGlySerIleIleGlyGlyThrPheLys 262
     ||||||||||||||::::::||||||||||||||||||||||||||||||
 704 TTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGTACTTTTAAG 753

 263 AlaIleLysGluArgSerSerThrProLysAlaProArgAspProArgLe 279
     ||||||:::||||||:::::::::|||||||||   ||||||||||||||
 754 GCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGACCCGCGCCT 803

 279 uProLysProLysGlyGlnThrValGlySerPheArgLysGlyLeuArgM 296
     ||||||||||:::|||||||||||||||||||||||||||||||||||||
 804 GCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGGGACTTCGAA 853

 296 etLeuProAspAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSer 312
     ||||||||:::|||||||||||||||||||||||||||:::|||||||||
 854 TGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTTAAGTTGTCT 903

 313 TrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTyrHisLeuTh 329
     ||||||||||||:::|||||||||   |||:::|||||||||:::|||||
 904 TGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATACAACTTAAC 953

 329 rTyrGluThrProGluGlyValValSerLeuGlnSerArgSerIleValM 346
     |||||||||||||:::|||:::||||||:::||||||:::|||:::||||
 954 ATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAAGTGTTGTAA 1003

 346 etThrValProSerTyrValAlaSerAsnIleLeuArgProLeuSerVal 362
     ||||||||||||||:::|||||||||::::::|||||||||||||||
1004 TGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCTCTTTCTGAA 1053

 363 AlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProValGlyAlaVa 379
     :::||||||:::|||||||||::::::|||||||||||||||:::|||||
1054 TCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGTTGCAGCAGT 1103

 379 lThrIleSerTyrProGlnGluAlaIleArgAspGluArgLeuValAspG 396
     |:::||||||||||||:::||||||||||||   |||   |||:::||||
1104 ATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTTTGATAGATG 1153

 396 lyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGlu 412
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1154 GTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAAGGAGTTGAA 1203

 413 ThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnArgAlaProLy 429
     ||||||||||||||||||||||||||||||||||||||||||||||||
1204 ACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCGCGCACCGCC 1253

 429 sGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluI 446
      ||||||:::||||||||||||||||||||||||:::   |||      |
1254 CGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAAACACCGGAA 1303

 446 leLeuSerLysThrGluSerGlnLeuValGluValValAspArgAspLeu 462
     |||||||||||:::|||::::::|||||||||:::|||||||||||||||
1304 TTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGACAGAGATTTG 1353

 463 ArgLysMetLeuIleLysProLysAlaGlnAspProLeuValValGlyVa 479
     |||||||||||||||||||||::::::   |||||||||   :::|||||
1354 AGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAAATTAGGAGT 1403

 479 lArgValTrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAspT 496
     |||||||||||||||||||||||||||||||||||||||||||:::|||
1404 TAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTCACTTTGATA 1453

 496 hrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeu 512
       |||:::|||||||||::::::::::::::::::|||   |||||||||
1454 TCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTACGAAGGGCTA 1503

 513 PheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyArgCysValGl 529
     |||||||||||||||||||||:::||||||||||||||||||||||||||
1504 TTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCGGTGTGTAGA 1553

 529 uGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuSerArgTyrA 546
     |||||||||||||:::|||   ||||||::::::|||:::||||||||||
1554 AGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGTCACGGTACG 1603

 546 laTyrLys 548
     ||||||||
1604 CTTACAAG 1611

seq_name: gb_pat2:AX084742
```

```
seq_documentation_block:
LOCUS       AX084742     1847 bp    DNA             PAT       09-MAR-2001
DEFINITION  Sequence 11 from Patent WO0112825.
ACCESSION   AX084742
VERSION     AX084742.1  GI:13274995
KEYWORDS    .
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
            Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE   1  (bases 1 to 1847)
  AUTHORS   Johnson,M.A., Volrath,S.L., Heifetz,P.B. and Law,M.D.
  TITLE     Herbicide-tolerant protoporphyrinogen oxidase
  JOURNAL   Patent: WO 0112825-A 11 22-FEB-2001;
            Syngenta Participations AG (CH)
FEATURES             Location/Qualifiers
     source          1. .1847
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
     CDS             55. .1686
                     /note="soybean protox-1"
                     /codon_start=1
                     /protein_id="CAC34013.1"
                     /db_xref="GI:13274996"
                     /translation="MVSVFNEILFPPNQTLLRPSLHSPTSFFTSPTRKFPRSRPNPIL
                     RCSIAEESTASPPKTRDSAPVDCVVVGGGVSGLCIAQALATKHANANVVVTEARDRVG
                     GNITTMERDGYLWEEGPNSFQPSDPMLTMVVDSGLKDELVLGDPDAPRFVLWNRKLRP
                     VPGKLTDLPFFDLMSIGGKIRAGFGALGIRPPPPGHEESVEEFVRRNLGDEVFERLIE
                     PFCSGVYAGDPSKLSMKAAFGKVWKLEKNGGSIIGGTFKAIQERNGASKPPRDPRLPK
                     PKGQTVGSFRKGLTMLPDAISARLGNKVKLSWKLSSISKLDSGEYSLTYETPEGVVSL
                     QCKTVVLTIPSYVASTLLRPLSAAAADALSKFYYPPVAAVSISYPKEAIRSECLIDGE
                     LKGFGQLHPRSQGVETLGTIYSSSLFPNRAPPGRVLLLNYIGGATNTGILSKTDSELV
                     ETVDRDLRKILINPNAQDPFVVGVRLWPQAIPQFLVGHLDLLDVAKASIRNTGFEGLF
                     LGGNYVSGVALGRCVEGAYEVAAEVNDFLTNRVYK"
BASE COUNT      473 a    416 c    453 g    505 t
ORIGIN


alignment_scores:
           Quality: 2062.50              Length:    550
             Ratio:  4.218                 Gaps:      7
 Percent Similarity:  88.909   Percent Identity: 73.273

alignment_block:
 US-09-508-418-2 x AX084742   ..

 Align seg 1/1  to: AX084742  from: 1  to: 1847

      3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
        ::::::|||:::         ||||||:::|||   |||   |||   ||
    102 TCTTCGCCCCTCCCTCCATTCCCCAACCTCTTTCTTCACCTCTCCCACTC 151

     19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerI 36
        |:::                              ||||||
    152 GAAA..............................ATTCCC.........T 162

     36 le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
           |||:::   |||   :::                     ||||||||
    163 CGCTCTCGCCCTAACCCTATT..................CTACGCTGCTC 194

     52 rValAlaLysAspTyrThrVal...ProSerSerAlaValAspGlyGlyP 68
        |:::|||::::::   |||:::   |||   ::::::   |||::::::|
    195 CATTGCGGAGGAATCCACCGCGTCTCCGCCCAAAACCAGAGACTCCGCCC 244

     68 roAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
        ||         :::|||||||||:::||||||:::|||:::|||||||||
    245 CC.........GTGGACTGCGTCGTCGTCGGCGGAGGCGTCAGCGGCCTC 285

     85 CysIleAlaGlnValMetSerAlaAsnTyr......ProAsnLeuMetVa 99
        ||||||||||||::::::::::::::::::         |||::::::||
    286 TGCATCGCCCAGGCCCTCGCCACCAAACACGCCAATGCCAACGTCGTCGT 335

  99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgA 116
     |||||||||||||||||||:::||||||||||||||||||:::|||||||
 336 CACGGAGGCCCGAGACCGCGTCGGCGGCAACATCACCACGATGGAGAGGG 385

 116 spGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 386 ACGGATACCTCTGGGAAGAAGGCCCCAACAGCTTCCAGCCTTCTGATCCA 435

 133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGl 149
     ||||||||||||:::||||||   ||||||||||||:::|||||||||||
 436 ATGCTCACCATGGTGGTGGACAGTGGTTTAAAGGATGAGCTTGTTTTGGG 485

 149 yAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
     |||||||:::|||||||||||||||||||||:::   |||||||||||||
 486 GGATCCTGATGCACCTCGGTTTGTGTTGTGGAACAGGAAGTTGAGGCCGG 535

 166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
     |||||:::|||||||||||||||||||||||||||||||||||||||
 536 TGCCCGGGAAGCTGACTGATTTGCCTTTCTTTGACTTGATGAGCATTGGT 585

 183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
     ||||||:::|||||||||||||||   :::|||:::||||||   |||||
 586 GGCAAAATCAGGGCTGGCTTTGGTGCGCTTGGAATTCGGCCTCCTCCTCC 635

 199 oGlyHisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyG 216
     ||||||||||||||||||||||:::|||||||||||||||||||||   |
 636 AGGTCATGAGGAATCGGTTGAAGAGTTTGTTCGTCGGAACCTTGGTGATG 685

 216 luValPheGluArgLeuIleGluProPheCysSerGlyValTyrValGly 232
     ||||||||||||||||||||||||||||||||||||||||||||:::|||
 686 AGGTTTTTGAACGGTTGATAGAGCCTTTTTGTTCAGGGGTCTATGCAGGC 735

 233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLe 249
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 736 GATCCTTCAAAATTAAGTATGAAAGCAGCATTCGGGAAAGTTTGGAAGCT 785

 249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
     ||||::::::||||||||||||||||||||||||||||||||||||:::|
 786 GGAAAAAAATGGTGGTAGCATTATTGGTGGAACTTTCAAAGCAATACAAG 835

 266 luArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
     |||||:::::::::   |||   |||||||||||||||||||||||||||
 836 AGAGAAATGGAGCTTCAAAACCACCTCGAGATCCGCGTCTGCCAAAACCA 885

 283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
     ||||||||||||||||||||||||||||||||||||   |||||||||||
 886 AAAGGTCAGACTGTTGGATCTTTCCGGAAGGGACTTACCATGTTGCCTGA 935

 299 pAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
     ||||||||||||||||||||||:::|||:::|||||||||||||||||||
 936 TGCAATTTCTGCCAGACTAGGCAACAAAGTAAAGTTATCTTGGAAGCTTT 985

 316 erSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThr 332
     ||||||||:::|||   ::::::|||   |||   |||||||||||||||
 986 CAAGTATTAGTAAACTGGATAGTGGAGAGTACAGTTTGACATATGAAACA 1035

 333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetThrValPr 349
     ||||||||||||||||||||||||   :::::::::|||:::|||:::||
1036 CCAGAAGGAGTGGTTTCTTTGCAGTGCAAAACTGTTGTCCTGACCATTCC 1085

 349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaA 366
     ||||||||||||||||::::::|||||||||||||||:::||||||||||
1086 TTCCTATGTTGCTAGTACATTGCTGCGTCCTCTGTCTGCTGCTGCTGCAG 1135

 366 spAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSer 382
     |||||||||||:::||||||||||||||||||:::||||||:::||||||
1136 ATGCACTTTCAAAGTTTTATTACCCTCCAGTTGCTGCAGTTTCCATATCC 1185

 383 TyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLy 399
     ||||||:::||||||||||||:::|||   |||:::||||||||||||||
1186 TATCCAAAAGAAGCTATTAGATCAGAATGCTTGATAGATGGTGAGTTGAA 1235
```

```
 399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyT 416
     |||||||||||||||||||||||||:::||||||||||||||||||||||
1236 GGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAACATTAGGAA 1285

 416 hrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
     ||||||||||||||||||||||||||||||||||||||   |||||||||
1286 CTATATACAGCTCATCACTATTCCCCAACCGAGCACCACCTGGAAGGGTT 1335

 433 LeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLy 449
     |||||||||||||||||||||||||||   |||      |||||||||||
1336 CTACTCTTGAATTACATTGGAGGAGCAACTAATACTGGAATTTTATCGAA 1385

 449 sThrGluSerGlnLeuValGluValValAspArgAspLeuArgLysMetL 466
     ||||:::|||:::|||||||||:::|||||||||||||||||||||:::|
1386 GACGGACAGTGAACTTGTGGAAACAGTTGATCGAGATTTGAGGAAAATCC 1435

 466 euIleLysProLysAlaGlnAspProLeuValValGlyValArgValTrp 482
     |||||:::|||:::||||||||||||:::|||||||||||||||:::|||
1436 TTATAAACCCAAATGCCCAGGATCCATTTGTAGTGGGGGTGAGACTGTGG 1485

 483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
     |||||||||||||||||||||||||||||||||||||||   |||:::::
1486 CCTCAAGCTATTCCACAGTTCTTAGTTGGCCATCTTGATCTTCTAGATGT 1535

 499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
     :|||||||||:::::::::::::::|||:::|||||||||||||||||||
1536 TGCTAAAGCTTCTATCAGAAATACTGGGTTTGAAGGGCTCTTCCTTGGGG 1585

 516 lyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1586 GTAATTATGTGTCTGGTGTTGCCTTGGGACGATGCGTTGAGGGAGCCTAT 1635

 533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
     |||||||||:::||||||:::   ||||||::::::   :::||||||
1636 GAGGTAGCAGCTGAAGTAAACGATTTTCTCACAAATAGAGTGTACAAA 1683

seq_name: gb_pat2:AX084750

seq_documentation_block:
LOCUS       AX084750     1784 bp    DNA             PAT       09-MAR-2001
DEFINITION  Sequence 19 from Patent WO0112825.
ACCESSION   AX084750
VERSION     AX084750.1  GI:13275003
KEYWORDS    .
SOURCE      rape.
  ORGANISM  Brassica napus
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
            Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 1784)
  AUTHORS   Johnson,M.A., Volrath,S.L., Heifetz,P.B. and Law,M.D.
  TITLE     Herbicide-tolerant protoporphyrinogen oxidase
  JOURNAL   Patent: WO 0112825-A 19 22-FEB-2001;
            Syngenta Participations AG (CH)
FEATURES             Location/Qualifiers
     source          1..1784
                     /organism="Brassica napus"
                     /db_xref="taxon:3708"
     CDS             47..1657
                     /note="Oilseed rape protox-1 coding sequence"
                     /codon_start=1
                     /protein_id="CAC34016.1"
                     /db_xref="GI:13275004"
                     /translation="MDLSLLRPQPFLSPFSNPFPRSRPYKPLNLRCSVSGGSVVGSST
                     IEGGGGGKTVTADCVIVGGGISGLCIAQALVTKHPDAAKNVMVTEAKDRVGGNIITRE
                     EQGFLWEEGPNSFQPSDPMLTMVVDSGLKDDLVLGDPTAPRFVLWNGKLRPVPSKLTD
                     LPFFDLMSIGGKIRAGFGAIGIRPSPPGREESVEEFVRRNLGDEVFERLIEPFCSGVY
                     AGDPAKLSMKAAFGKVWKLEENGGSIIGGAFKAIQAKNKAPKTTRDPRLPKPKGQTVG
                     SFRKGLTMLPEAISARLGDKVKVSWKLSSITKLASGEYSLTYETPEGIVTVQSKSVVM
                     TVPSHVASSLLRPLSDSAAEALSKLYYPPVAAVSISYAKEAIRSECLIDGELKGFGQL
                     HPRTQKVETLGTIYSSSLFPNRAPPGRVLLLNYIGGATNTGILSKSEGELVEAVDRDL
                     RKMLIKPSSTDPLVLGVKLWPQAIPQFLIGHIDLVDAAKASLSSSGHEGLFLGGNYVA
                     GVALGRCVEGAYETATQVNDFMSRYAYK"
```

```
BASE COUNT      464 a    379 c    478 g    463 t
ORIGIN

alignment_scores:
            Quality: 2006.50             Length:    535
              Ratio:  4.189                Gaps:      4
 Percent Similarity:  89.533   Percent Identity: 73.084

alignment_block:
 US-09-508-418-2 x AX084750   ..

 Align seg 1/1  to: AX084750  from: 1  to: 1784

      21 ProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerIleSe 37
         |||      ||||||:::               |||||||||:::
      68 CCGCAGCCATTCCTATCG...............CCATTCTCAAATCCATT 102

      37 rLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSerValA 54
             |||   |||         :::      :::   ||||||||||||:
     103 TCCTCGG...TCGCGTCCCTACAAGCCTCTCAACCTCCGTTGCTCCGTAT 149

      54 laLysAspTyrThrValProSerSerAlaValAspGlyGlyProAlaAla 70
         ::         :::|||   |||||||:::::::::||||||   ::::::
     150 CCGGTGGATCCGTCGTCGGCTCTTCTACAATCGAAGGCGGAGGAGGAGGT 199

      71 Glu.........LeuAspCysValIleValGlyAlaGlyIleSerGlyLe 84
         :::            ||||||||||||||||||:::||||||||||||||
     200 AAAACCGTCACGGCGGACTGCGTGATCGTCGGCGGAGGAATCAGCGGCCT 249

      84 uCysIleAlaGlnValMetSerAlaAsnTyrPro............AsnL 97
         |||||||||||||::::::   :::::::::|||            |||:
     250 GTGCATTGCGCAAGCGCTCGTGACGAAGCACCCAGACGCTGCAAAGAATG 299

      97 euMetValThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrVal 113
         ::|||||||||||||||:::||||||:::||||||||||||   |||
     300 TGATGGTGACGGAGGCGAAGGACCGTGTGGGAGGGAATATCATCACGCGA 349

     114 GluArgAspGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSe 130
         |||::::::|||:::|||||||||||||||||||||||||||||||||||
     350 GAGGAGCAAGGGTTTCTATGGGAAGAAGGTCCCAATAGCTTTCAGCCGTC 399

     130 rAspProMetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuV 147
         |||||||||||||||||||:::||||||   |||||||||||||||||||
     400 TGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGAAAGATGATCTAG 449

     147 alLeuGlyAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeu 163
         ||||||||||||||:::|||||||||||||||||||||:::|||||||||
     450 TCTTGGGAGATCCTACTGCTCCGAGGTTTGTGTTGTGGAATGGGAAGCTG 499

     164 ArgProValProSerLysLeuThrAspLeuProPhePheAspLeuMetSe 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     500 AGGCCGGTTCCGTCGAAGCTAACTGACTTGCCTTTCTTTGACTTGATGAG 549

     180 rIleProGlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProS 197
         ||||   ||||||:::|||||||||||||||   ||||||:::|||||||
     550 TATTGGAGGGAAGATTAGAGCTGGGTTTGGTGCCATTGGTATTCGACCTT 599

     197 erProProGlyHisGluGluSerValGluGlnPheValArgArgAsnLeu 213
         |||||||||||:::|||||||||||||||:::||||||||||||||||||
     600 CACCTCCGGGTCGTGAGGAATCAGTGGAAGAGTTTGTAAGGCGTAATCTT 649

     214 GlyGlyGluValPheGluArgLeuIleGluProPheCysSerGlyValTy 230
         |||   ||||||||||||||||||||||||||||||||||||||||||||
     650 GGTGATGAGGTTTTTGAGCGCTTGATTGAACCCTTTTGCTCAGGTGTTTA 699

     230 rValGlyAspProSerLysLeuSerMetLysAlaAlaPheGlyLysValT 247
         |:::|||||||||:::||||||||||||||||||||||||||||||||||
     700 TGCGGGAGATCCTGCGAAACTGAGTATGAAAGCAGCTTTTGGGAAGGTTT 749

     247 rpLysLeuGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAla 263
         ||||||||||||||:::|||||||||||||||||||||:::|||||||||
     750 GGAAGCTAGAGGAGAATGGTGGGAGCATCATTGGTGGTGCTTTTAAGGCA 799
```

```
 264 IleLysGluArgSerSerThrProLysAlaProArgAspProArgLeuPr 280
     |||:::   ::::::::::::||||||:::   |||||||||||||||||
 800 ATTCAAGCGAAAAATAAAGCTCCCAAGACAACCCGAGATCCGCGTCTGCC 849

 280 oLysProLysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetL 297
     |||||||||||||||||||||||||||||||||||||||||||   ||||
 850 AAAGCCAAAGGGCCAAACTGTTGGTTCTTTCAGGAAAGGACTCACAATGC 899

 297 euProAspAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrp 313
     |||||:::|||||||||||||||||||||:::|||:::|||:::||||||
 900 TGCCAGAGGCAATCTCCGCAAGGTTGGGTGACAAGGTGAAAGTTTCTTGG 949

 314 LysLeuSerSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTy 330
     |||||||||||||||||||||      :::|||   |||   ||||||||
 950 AAGCTCTCAAGTATCACTAAGCTGGCCAGCGGAGAATATAGCTTAACTTA 999

 330 rGluThrProGluGlyValValSerLeuGlnSerArgSerIleValMetT 347
     ||||||||||||||||:::|||::::::||||||:::|||:::|||||||
1000 CGAAACTCCGGAGGGTATAGTCACTGTACAGAGCAAAAGTGTAGTGATGA 1049

 347 hrValProSerTyrValAlaSerAsnIleLeuArgProLeuSerValAla 363
     |||||||||||:::|||||||||::::::|||||||||||||||   :::
1050 CTGTGCCATCTCATGTTGCTAGTAGTCTCTTGCGCCCTCTCTCTGATTCT 1099

 364 AlaAlaAspAlaLeuSerAsnPheTyrTyrProProValGlyAlaValTh 380
     ||||||:::|||||||||::::::|||||||||||||||:::||||||::
1100 GCAGCTGAAGCGCTCTCAAAACTCTACTATCCGCCAGTTGCAGCCGTATC 1149

 380 rIleSerTyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyG 397
     :|||||||||   :::||||||||||||:::|||   |||:::|||||||
1150 CATCTCATACGCGAAAGAAGCAATCCGAAGCGAATGCTTAATAGATGGTG 1199

 397 luLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThr 413
     ||||||||||||||||||||||||||||||||||||||   |||||||||
1200 AACTAAAAGGGTTCGGCCAGTTGCATCCACGCACGCAAAAAGTGGAAACT 1249

 414 LeuGlyThrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGl 430
     |||||||||||||||||||||||||||||||||||||||||||||   ||
1250 CTTGGAACAATATACAGTTCATCGCTCTTTCCCAACCGAGCACCGCCTGG 1299

 430 yArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleL 447
     ||||||||||||||||||||||||||||||||||   |||      ||||
1300 AAGAGTATTGCTATTGAACTACATCGGTGGAGCTACCAACACTGGGATCT 1349

 447 euSerLysThrGluSerGlnLeuValGluValValAspArgAspLeuArg 463
     ||||||||:::|||::::::|||||||||:::||||||||||||||||||
1350 TATCAAAGTCGGAAGGTGAGTTAGTGGAAGCAGTAGATAGAGACTTGAGG 1399

 464 LysMetLeuIleLysProLysAlaGlnAspProLeuValValGlyValAr 480
     ||||||||||||||||||::::::   ||||||||||||:::||||||::
1400 AAGATGCTGATAAAGCCAAGCTCGACCGATCCACTTGTACTTGGAGTAAA 1449

 480 gValTrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrL 497
     ::::|||||||||||||||||||||||||||:::||||||:::|||   :
1450 ATTATGGCCTCAAGCCATTCCTCAGTTTCTGATAGGTCACATTGATTTGG 1499

 497 euSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPhe 513
     ::::::::|||||||||:::::::::::::::|||   ||||||||||||
1500 TAGACGCAGCGAAAGCATCGCTCTCGTCATCTGGTCATGAGGGCTTATTC 1549

 514 LeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGl 530
     ||||||||||||||||||:::|||||||||||||||||||||||||||||
1550 TTGGGTGGAAATTACGTTGCCGGTGTAGCATTGGGTCGGTGTGTGGAAGG 1599

 530 yAlaTyrGluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaT 547
     ||||||||||:::|||::::::|||:::   |||:::|||||||||||||
1600 TGCTTATGAAACTGCAACCCAAGTGAATGATTTCATGTCAAGGTATGCTT 1649

 547 yrLys 548
     |||||
1650 ACAAG 1654
```

```
seq_name: gb_pl2:AF160961

seq_documentation_block:
LOCUS       AF160961     1925 bp    mRNA            PLN       01-OCT-1999
DEFINITION  Cichorium intybus protoporphyrinogen IX oxidase (PPX1) mRNA,
            complete cds; nuclear gene for plastid product.
ACCESSION   AF160961
VERSION     AF160961.1  GI:6002911
KEYWORDS    .
SOURCE      chicory.
  ORGANISM  Cichorium intybus
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            II; Asterales; Asteraceae; Lactuceae; Cichorium.
REFERENCE   1  (bases 1 to 1925)
  AUTHORS   Adomat,C. and Boeger,P.
  TITLE     Cloning, Sequence, Expression and Characterization of
            Protoporphyrinogen IX Oxidase from Chicory
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1925)
  AUTHORS   Adomat,C. and Boeger,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-1999) Plant Biochemistry, University of Konstanz,
            Konstanz 78434, Germany
FEATURES             Location/Qualifiers
     source          1..1925
                     /organism="Cichorium intybus"
                     /db_xref="taxon:13427"
                     /note="variety: foliosum"
     gene            1..1925
                     /gene="PPX1"
     CDS             91..1758
                     /gene="PPX1"
                     /EC_number="1.3.3.4"
                     /codon_start=1
                     /product="protoporphyrinogen IX oxidase"
                     /protein_id="AAF00194.1"
                     /db_xref="GI:6002912"
                     /translation="MTSLTDVCSLNCCRSWSSLPPPVSGGSLTSKNPRYLITYSPAHR
                     KCNRWRFRCSIAKDSPITPPISNEFNSQPLLDCVIVGAGISGLCIAQALATKHASVSP
                     DVIVTEARDRVGGNISTVERDGYLWEEGPNSFQPSDAMLTMVVDSGLKDDLVLGDPTA
                     PRFVLWGGDLKPVPSKPADLPFFDLMSFPGKLRAGFGALGFRPSPPDREESVEEFVRR
                     NLGDEVFERLIEPFCSGVYAGDPSKLSMKAAFGKVWNLEQNGGSIVGGAFKAIQDRKN
                     SQKPPRDPRLPKPKGQTVGSFRKGQAMLPNAISTRLGSRVKLCWKLTSISKLENRGYN
                     LTYETPQGFESLQTKTIVMTVPSYVASDLLRPLSLGAADALSKFYYPPVAAVSISYPK
                     DAIRADRLIDGQLKGFGQLHPRSQGVETLGTIYSSSLFPNRAPPGRVLLLNYIGGATN
                     PEILSKTEGEIVDAVDRDLRTMLIRRDAEDPLTLGVRVWPRAIPQFLIGHYDILDSAK
                     AALSSGGFQGMFLGGNYVSGVALGKCVEAAYDVAAEVMNFLSQGVYK"
BASE COUNT      507 a    398 c    496 g    524 t
ORIGIN


alignment_scores:
            Quality: 2004.50          Length:    536
              Ratio:  4.194             Gaps:      3
 Percent Similarity: 89.179  Percent Identity: 71.455

alignment_block:
 US-09-508-418-2 x AF160961   ..

 Align seg 1/1  to: AF160961  from: 1  to: 1925

      18 SerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSe 34
         |||::::::   |||:::      |||    .    :::|||   :::||
     160 TCTGGTGGGTCGTTGACGTCAAAGAATCCTAGGTACCTAATCACGTATAG 209

      34 rSerIleSerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgC 51
         |         ::::::            ||||||   ||||||   ||||
     210 TCCGGCGCATCGCAAA............TGCAATAGGTGGAGGTTCCGCT 247

      51 ysSerValAlaLysAspTyrThrVal...ProSerSerAlaValAspGly 66
         |||||:::|||||||||      :::   |||      :::   :::
     248 GCTCTATAGCCAAGGATTCCCCAATTACTCCTCCCATTTCAAATGAGTTC 297
```

```
  67 GlyProAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGl 83
     :::            |||||||||||||||||||||||||||||||||||
 298 AACTCTCAGCCATTGTTGGACTGTGTCATTGTGGGCGCCGGCATTAGCGG 347

  83 yLeuCysIleAlaGlnValMetSerAlaAsnTyr............ProA 96
     ||||||||||||||||::::::::::::::::::            |||:
 348 CCTTTGCATTGCGCAGGCCCTAGCGACTAAACACGCCTCCGTCTCTCCGG 397

  96 snLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThr 112
     ::::::::|||||||||||||||||||||:::||||||||||||:::|||
 398 ATGTGATCGTCACCGAGGCACGAGACAGAGTCGGGGTAATATATCAACG 447

 113 ValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnPr 129
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 448 GTTGAAAGGGATGGCTATCTCTGGGAAGAAGGTCCTAACAGCTTCCAGCC 497

 129 oSerAspProMetLeuThrMetAlaValAspCysGlyLeuLysAspAspL 146
     |||||||   ||||||||||||:::||||||   ||||||||||||||||
 498 ATCTGATGCCATGCTCACCATGGTGGTGGATAGTGGGTTGAAGGATGATT 547

 146 euValLeuGlyAspProAsnAlaProArgPheValLeuTrpLysGlyLys 162
     |||||||||||||||||:::|||||||||||||||||||||   |||
 548 TGGTGTTAGGTGACCCAACAGCACCCCGCTTTGTATTATGGGGAGGTGAT 597

 163 LeuArgProValProSerLysLeuThrAspLeuProPhePheAspLeuMe 179
     |||:::|||||||||||||||   :::|||||||||||||||||||||||
 598 TTGAAACCGGTTCCTTCCAAACCGGCTGACCTCCCTTTCTTTGACCTCAT 647

 179 tSerIleProGlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgP 196
     ||||:::|||||||||||||||||||||||||||   :::|||:::||||
 648 GAGCTTTCCTGGAAAACTCAGAGCCGGTTTTGGTGCTCTTGGATTCCGTC 697

 196 roSerProProGlyHisGluGluSerValGluGlnPheValArgArgAsn 212
     |||||||||||   :::|||||||||||||||:::|||||||||||||||
 698 CTTCACCTCCAGATCGCGAAGAATCGGTTGAGGAGTTTGTTAGACGTAAT 747

 213 LeuGlyGlyGluValPheGluArgLeuIleGluProPheCysSerGlyVa 229
     ||||||   |||||||||||||||||||||||||||||||||||||||||
 748 CTTGGAGATGAAGTTTTCGAACGCTTGATAGAACCTTTTTGCTCAGGTGT 797

 229 lTyrValGlyAspProSerLysLeuSerMetLysAlaAlaPheGlyLysV 246
     ||||:::|||||||||||||||||||||||||||||||||||||||||||
 798 TTATGCTGGTGATCCATCAAAACTTAGTATGAAAGCAGCATTTGGGAAGG 847

 246 alTrpLysLeuGluGluThrGlyGlySerIleIleGlyGlyThrPheLys 262
     |||||:::||||||::::::||||||||||||:::||||||:::||||||
 848 TCTGGAATCTGGAGCAAAATGGTGGTAGCATTGTTGGTGGAGCCTTCAAG 897

 263 AlaIleLysGluArgSerSerThrProLysAlaProArgAspProArgLe 279
     ||||||::::::|||:::::::::   |||   |||||||||||||||||
 898 GCTATTCAGGACAGAAAGAATAGTCAAAAGCCTCCACGGGACCCGAGGTT 947

 279 uProLysProLysGlyGlnThrValGlySerPheArgLysGlyLeuArgM 296
     |||||||||||||||||||||||||||||||||||||||||||      |
 948 ACCGAAACCAAAGGGCCAAACTGTTGGATCTTTTAGGAAAGGACAAGCGA 997

 296 etLeuProAspAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSer 312
     ||||||||:::|||||||||:::||||||||||||::::::||||||
 998 TGTTGCCTAATGCAATCTCAACGAGGTTAGGTAGCAGAGTGAAATTGTGT 1047

 313 TrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTyrHisLeuTh 329
     |||||||||:::||||||:::|||   |||:::   ||||||:::|||||
1048 TGGAAGCTCACGAGTATTTCAAAATTGGAGAATAGAGGTTATAATTTGAC 1097

 329 rTyrGluThrProGluGlyValValSerLeuGlnSerArgSerIleValM 346
     |||||||||||||:::|||      |||||||||:::::::::|||||||
1098 ATATGAAACACCACAAGGATTTGAAAGTCTGCAGACTAAAACTATCGTGA 1147

 346 etThrValProSerTyrValAlaSerAsnIleLeuArgProLeuSerVal 362
     ||||||||||||||||||||||||||::::::|||||||||||||||:::
1148 TGACTGTTCCATCCTACGTGGCGAGTGACTTGTTGCGTCCGCTTTCGTTG 1197

 363 AlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProValGlyAlaVa 379
     :::||||||||||||||||||:::||||||||||||||||||:::|||||
1198 GGTGCAGCAGATGCATTGTCAAAATTTTATTATCCTCCGGTTGCAGCTGT 1247

 379 lThrIleSerTyrProGlnGluAlaIleArgAspGluArgLeuValAspG 396
     |:::||||||||||||::::::|||||||||   :::||||||:::||||
1248 ATCAATTTCATATCCAAAAGACGCAATTCGTGCTGACCGGCTGATTGATG 1297

 396 lyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGlu 412
     ||:::||||||||||||||||||||||||||||||:::||||||||||||
1298 GTCAACTCAAAGGTTTTGGGCAATTGCATCCACGAAGTCAAGGGGTGGAA 1347

 413 ThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnArgAlaProLy 429
     ||||||||||||||||||||||||||||||||||||||||||||||||
1348 ACTTTAGGTACGATCTACAGTTCATCTCTTTTCCCTAACCGAGCGCCACC 1397

 429 sGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluI 446
      ||||||||||||||||||||||||||||||||||||   ||||||||||
1398 TGGAAGGGTTCTGCTCTTGAACTACATCGGAGGGGCTACAAATCCTGAAA 1447

 446 leLeuSerLysThrGluSerGlnLeuValGluValValAspArgAspLeu 462
     |||||||||||||||||:::::::::|||::::::|||||||||||||||
1448 TTCTATCAAAGACGGAGGGCGAAATTGTGGATGCGGTGGACCGGGACCTA 1497

 463 ArgLysMetLeuIleLysProLysAlaGlnAspProLeuValValGlyVa 479
     |||   |||||||||:::      |||:::|||||||||::::::|||||
1498 CGGACGATGCTGATAAGGCGTGATGCGGAAGATCCATTGACGTTGGGGGT 1547

 479 lArgValTrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAspT 496
     |||||||||||||:::||||||||||||||||||:::||||||   |||
1548 GCGGGTGTGGCCTCGAGCAATCCCGCAGTTTCTGATCGGTCATTATGACA 1597

 496 hrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeu 512
       |||::::::||||||||||||::::::::::::|||::::::|||:::
1598 TTCTAGATTCTGCAAAAGCTGCTCTGAGTAGCGGTGGATTCCAAGGTATG 1647

 513 PheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyArgCysValGl 529
     |||||||||||||||||||||||||||||||||||||||:::||||||||
1648 TTTCTTGGTGGCAACTATGTGTCTGGTGTGGCTTTAGGTAAATGTGTCGA 1697

 529 uGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuSerArgTyrA 546
     |:::||||||:::||||||:::||||||   :::|||||||||:::   :
1698 GGCTGCTTATGATGTTGCCGCTGAGGTAATGAACTTTTTGTCGCAAGGGG 1747

 546 laTyrLys 548
     ::||||||
1748 TGTACAAG 1755

seq_name: gb_pl2:AF218052

seq_documentation_block:
LOCUS       AF218052     1608 bp    mRNA            PLN       15-FEB-2000
DEFINITION  Zea mays protoporphyrinogen IX oxidase mRNA, complete cds.
ACCESSION   AF218052
VERSION     AF218052.1  GI:6715440
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
            Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1608)
  AUTHORS   Volrath,S.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-DEC-1999) Novartis Agricultural Biotechnology, 3054
            Cornwallis Road, Research Triangle Park, NC 27709, USA
FEATURES             Location/Qualifiers
     source          1. .1608
                     /organism="Zea mays"
                     /cultivar="B73 inbred"
                     /db_xref="taxon:4577"
     CDS             1. .1608
                     /codon_start=1
```

```
                     /product="protoporphyrinogen IX oxidase"
                     /protein_id="AAF26417.1"
                     /db_xref="GI:6715441"
                     /translation="MVAATATAMATAASPLLNGTRIPARLRHRGLSVRCAAVAGGAAE
                     APASTGARLSADCVVVGGGISGLCTAQALATRHGVGDVLVTEARARPGGNITTVERPE
                     EGYLWEEGPNSFQPSDPVLTMAVDSGLKDDLVFGDPNAPRFVLWEGKLRPVPSKPADL
                     PFFDLMSIPGKLRAGLGALGIRPPPPGREESVEEFVRRNLGAEVFERLIEPFCSGVYA
                     GDPSKLSMKAAFGKVWRLEETGGSIIGGTIKTIQERSKNPKPPRDARLPKPKGQTVAS
                     FRKGLAMLPNAITSSLGSKVKLSWKLTSITKSDDKGYVLEYETPEGVVSVQAKSVIMT
                     IPSYVASNILRPLSSDAADALSRFYYPPVAAVTVSYPKEAIRKECLIDGELQGFGQLH
                     PRSQGVETLGTIYSSSLFPNRAPDGRVLLLNYIGGATNTGIVSKTESELVEAVDRDLR
                     KMLINSTAVDPLVLGVRVWPQAIPQFLVGHLDLLEAAKAALDRGGYDGLFLGGNYVAG
                     VALGRCVEGAYESASQISDFLTKYAYK"
BASE COUNT     360 a    430 c    464 g    354 t
ORIGIN

alignment_scores:
           Quality: 1989.50             Length:    545
             Ratio:  4.260                Gaps:      6
 Percent Similarity:  85.688  Percent Identity: 72.294

alignment_block:
 US-09-508-418-2 x AF218052   ..

 Align seg 1/1  to: AF218052  from: 1  to: 1608

    17 SerSerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPh 33
       :::::::::|||||||||      ||||||   |||      ||||||
    31 ACCGCTGCATCGCCGCTA......CTCAACGGGACCCGA...ATACCTGC 71

    33 eSerSerIleSerLysArgAsnSerValAsnCysAsnGlyTrpArgThrA 50
                    :::      ||||||:::|||
    72 GCGGCTCCGCCATCGAGGACTCAGCGTGCGCTGC................ 105

    50 rgCysSerValAlaLysAspTyrThrValProSerSerAlaValAspGly 66
                                          :::||||||   |||
   106 ..................................GCTGCTGTGGCGGGC 120

    67 GlyProAla...........................AlaGluLeuAspCy 74
       |||   |||                              :::   |||||
   121 GGCGCGGCCGAGGCACCGGCATCCACCGGCGCGCGGCTGTCCGCGGACTG 170

    74 sValIleValGlyAlaGlyIleSerGlyLeuCysIleAlaGlnValMetS 91
       ||||:::||||||:::||||||||||||||||||   ||||||:::::::
   171 CGTCGTGGTGGGCGGAGGCATCAGTGGCCTCTGCACCGCGCAGGCGCTGG 220

    91 erAlaAsnTyr......ProAsnLeuMetValThrGluAlaArgAspArg 105
       :::::::::::         :::::::::|||||||||||||||   |||
   221 CCACGCGGCACGGCGTCGGGGACGTGCTTGTCACGGAGGCCCGCGCCCGC 270

   106 AlaGlyGlyAsnIleThrThrValGluArg......AspGlyTyrLeuTr 120
          |||||||||||||||||||||||||||      :::|||||||||||
   271 CCCGGCGGCAACATTACCACCGTCGAGCGCCCCGAGGAAGGGTACCTCTG 320

   120 pGluGluGlyProAsnSerPheGlnProSerAspProMetLeuThrMetA 137
       |||||||||||||||||||||||||||||||||||||:::||||||||||
   321 GGAGGAGGGTCCCAACAGCTTCCAGCCCTCCGACCCCGTTCTCACCATGG 370

   137 laValAspCysGlyLeuLysAspAspLeuValLeuGlyAspProAsnAla 153
       ||||||||   |||||||||||||||||||||:::|||||||||||||||
   371 CCGTGGACAGCGGACTGAAGGATGACTTGGTTTTGGGGACCCAAACGCG 420

   154 ProArgPheValLeuTrpLysGlyLysLeuArgProValProSerLysLe 170
       ||||||||||||||||||:::|||||||||||||||||||||||||||
   421 CCGCGTTTCGTGCTGTGGAGGGGAAGCTGAGGCCCGTGCCATCCAAGCC 470

   170 uThrAspLeuProPhePheAspLeuMetSerIleProGlyLysLeuArgA 187
        :::||||||||||||||||||||||||||||||||||||||||||||||
   471 CGCCGACCTCCCGTTCTTCGATCTCATGAGCATCCCAGGGAAGCTCAGGG 520

   187 laGlyPheGlyProIleGlyLeuArgProSerProProGlyHisGluGlu 203
       |||||:::|||   :::|||:::||||||   |||||||||:::||||||
```

```
   521 CCGGTCTAGGCGCGCTTGGCATCCGCCCGCCTCCTCCAGGCCGCGAAGAG 570

   204 SerValGluGlnPheValArgArgAsnLeuGlyGlyGluValPheGluAr 220
       |||||||||:::|||||||||||||||||||||:::||||||||||||||
   571 TCAGTGGAGGAGTTCGTGCGCCGCAACCTCGGTGCTGAGGTCTTTGAGCG 620

   220 gLeuIleGluProPheCysSerGlyValTyrValGlyAspProSerLysL 237
       |||||||||||||||||||||||||||||||:::||||||||||||||||
   621 CCTCATTGAGCCTTTCTGCTCAGGTGTCTATGCTGGTGATCCTTCTAAGC 670

   237 euSerMetLysAlaAlaPheGlyLysValTrpLysLeuGluGluThrGly 253
       ||||||||||||||||||||||||||||||||:::|||||||||||||||
   671 TCAGCATGAAGGCTGCATTTGGGAAGGTTTGGCGGTTGGAAGAAACTGGA 720

   254 GlySerIleIleGlyGlyThrPheLysAlaIleLysGluArgSerSerTh 270
       |||||||||||||||||||||:::|||:::|||:::|||||||||:::::
   721 GGTAGTATTATTGGTGGAACCATCAAGACAATTCAGGAGAGGAGCAAGAA 770

   270 rProLysAlaProArgAspProArgLeuProLysProLysGlyGlnThrV 287
       :||||||   |||||||||   ||||||||||||||||||||||||||||
   771 TCCAAAACCACCGAGGGATGCCCGCCTTCCGAAGCCAAAAGGGCAGACAG 820

   287 alGlySerPheArgLysGlyLeuArgMetLeuProAspAlaIleSerAla 303
       ||:::||||||||||||||||||   |||||||||:::||||||::::::
   821 TTGCATCTTTCAGGAAGGGTCTTGCCATGCTTCCAAATGCCATTACATCC 870

   304 ArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSerSerIleThrLy 320
          ||||||||||||:::||||||||||||||||||:::|||||||||||
   871 AGCTTGGGTAGTAAAGTCAAACTATCATGGAAACTCACGAGCATTACAAA 920

   320 sSerGluLysGlyGlyTyrHisLeuThrTyrGluThrProGluGlyValV 337
       ||||:::      ||||||   |||   ||||||||||||||||||||||
   921 ATCAGATGACAAGGGATATGTTTTGGAGTATGAAACGCCAGAAGGGGTTG 970

   337 alSerLeuGlnSerArgSerIleValMetThrValProSerTyrValAla 353
       |||||:::|||::::::|||::::::||||||:::|||||||||||||||
   971 TTTCGGTGCAGGCTAAAAGTGTTATCATGACTATTCCATCATATGTTGCT 1020

   354 SerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAlaLeuSerAs 370
       ||||||||||||||||||||||||      ||||||||||||||||||::
  1021 AGCAACATTTTGCGTCCACTTTCAAGCGATGCTGCAGATGCTCTATCAAG 1070

   370 nPheTyrTyrProProValGlyAlaValThrIleSerTyrProGlnGluA 387
       :||||||||||||||||||:::|||||||||:::|||||||||:::||||
  1071 ATTCTATTATCCACCGGTTGCTGCTGTAACTGTTTCGTATCCAAAGGAAG 1120

   387 laIleArgAspGluArgLeuValAspGlyGluLeuLysGlyPheGlyGln 403
       ||||||||   |||   ||||:::||||||||||||:::||||||||||||
  1121 CAATTAGAAAAGAATGCTTAATTGATGGGGAACTCCAGGGCTTTGGCCAG 1170

   404 LeuHisProArgThrGlnGlyValGluThrLeuGlyThrIleTyrSerSe 420
       ||||||||||||:::|||||||||||||||||||||||||||||||||||
  1171 TTGCATCCACGTAGTCAAGGAGTTGAGACATTAGGAACAATATACAGTTC 1220

   420 rSerLeuPheProAsnArgAlaProLysGlyArgValLeuLeuLeuAsnT 437
       |||||||||||||||||||||||||   ||||||||||||||||||||||
  1221 CTCACTCTTTCCAAATCGTGCTCCTGACGGTAGGGTGTTACTTCTAAACT 1270

   437 yrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThrGluSerGln 453
       ||||||||||||||   |||      ||||:::|||||||||||||||:::
  1271 ACATAGGAGGTGCTACAAACACAGGAATTGTTTCCAAGACTGAAAGTGAG 1320

   454 LeuValGluValValAspArgAspLeuArgLysMetLeuIleLysProLy 470
       |||||||||:::||||||||||||||||||||||||||||||:::
  1321 CTGGTCGAAGCAGTTGACCGTGACCTCCGAAAAATGCTTATAAATTCTAC 1370

   470 sAlaGlnAspProLeuValValGlyValArgValTrpProGlnAlaIleP 487
        |||   ||||||||||||:::||||||||||||||||||||||||||||
  1371 AGCAGTGGACCCTTTAGTCCTTGGTGTTCGAGTTTGGCCACAAGCCATAC 1420

   487 roGlnPheLeuValGlyHisLeuAspThrLeuSerThrAlaLysAlaAla 503
       ||||||||||||||||||||||||||   |||::::::||||||||||||
  1421 CTCAGTTCCTGGTAGGACATCTTGATCTTCTGGAAGCCGCAAAAGCTGCC 1470
```

```
     504 MetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAsnTyrValSe 520
         ::::::   :::|||   :::|||||||||||||||||||||||||||::
    1471 CTGGACCGAGGTGGCTACGATGGGCTGTTCCTAGGAGGGAACTATGTTGC 1520

     520 rGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluValAlaSerG 537
         :|||||||||||||||||||||||||||||||||||||||   ||||||:
    1521 AGGAGTTGCCCTGGGCAGATGCGTTGAGGGCGCGTATGAAAGTGCCTCGC 1570

     537 luValThrGlyPheLeuSerArgTyrAlaTyrLys 548
         ::::::::   ||||||::::::||||||||||||
    1571 AAATATCTGACTTCTTGACCAAGTATGCCTACAAG 1605

seq_name: gb_pat2:AX084740

seq_documentation_block:
LOCUS       AX084740     1811 bp    DNA             PAT       09-MAR-2001
DEFINITION  Sequence 9 from Patent WO0112825.
ACCESSION   AX084740
VERSION     AX084740.1  GI:13274993
KEYWORDS    .
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae;
            Triticum.
REFERENCE   1  (bases 1 to 1811)
  AUTHORS   Johnson,M.A., Volrath,S.L., Heifetz,P.B. and Law,M.D.
  TITLE     Herbicide-tolerant protoporphyrinogen oxidase
  JOURNAL   Patent: WO 0112825-A 9 22-FEB-2001;
            Syngenta Participations AG (CH)
FEATURES             Location/Qualifiers
     source          1..1811
                     /organism="Triticum aestivum"
                     /db_xref="taxon:4565"
     CDS             <3..1589
                     /note="wheat protox-1"
                     /codon_start=1
                     /protein_id="CAC34012.1"
                     /db_xref="GI:13274994"
                     /translation="ATMATATVAAASPLRGRVTGRPHRVRPRCATASSATETPAAPGV
                     RLSAECVIVGAGISGLCTAQALATRYGVSDLLVTEARDRPGGNITTVERPDEGYLWEE
                     GPNSFQPSDPVLTMAVDSGLKDDLVFGDPNAPRFVLWEGKLRPVPSKPGDLPFFSLMS
                     IPGKLRAGLGALGIRPPPPGREESVEEFVRRNLGAEVFERLIEPFCSGVYAGDPSKLS
                     MKAAFGKVWRLEEIGGSIIGGTIKAIQDKGKNPKPPRDPRLPAPKGQTVASFRKGLAM
                     LPNAIASRLGSKVKLSWKLTSITKADNQGYVLGYETPEGLVSVQAKSVIMTIPSYVAS
                     DILRPLSIDAADALSKFYYPPVAAVTVSYPKEAIRKECLIDGELQGFGQLHPRSQGVE
                     TLGTIYSSSLFPNRAPAGRVLLLNYIGGSTNTGIVSKTESDLVGAVDRDLRKMLINPR
                     AADPLALGVRVWPQAIPQFLIGHLDRLAAAKSALGQGGYDGLFLGGNYVAGVALGRCI
                     EGAYESASQVSDFLTKYAYK"
BASE COUNT      423 a    463 c    517 g    408 t
ORIGIN


alignment_scores:
            Quality: 1987.00             Length:    505
              Ratio:  4.310                Gaps:      3
 Percent Similarity: 91.287   Percent Identity: 75.050

alignment_block:
 US-09-508-418-2 x AX084740   ..

 Align seg 1/1  to: AX084740  from: 1  to: 1811

      48 ArgThrArgCysSerValAlaLysAspTyrThrValProSerSerAlaVa 64
         |||   ||||||::::::|||::::::   |||         :::|||
      78 CGCCCGCGTTGCGCTACCGCGAGCAGCGCGACCGAGACTCCGGCGGCG.. 125

      64 lAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaGlyI 81
                |||         :::   :::||||||||||||||||||||||
     126 ....CCCGGCGTGCGGCTGTCCGCGGAATGCGTCATTGTGGGCGCCGGCA 171

      81 leSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr......Pro 95
         ||||||||||||||   ||||||:::::::::::::::|||
     172 TCAGCGGCCTCTGCACCGCGCAGGCGCTGGCCACCCGATACGGCGTCAGC 221

      96 AsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsnIleThrTh 112
         :::|||:::|||||||||||||||||||||   |||||||||||||||||
     222 GACCTGCTCGTCACGGAGGCCCGCGACCGCCCGGGCGGCAACATCACCAC 271

     112 rValGluArg......AspGlyTyrLeuTrpGluGluGlyProAsnSerP 127
         ||||||||||      :::|||||||||||||||||||||||||||||||
     272 CGTCGAGCGTCCCGACGAGGGGTACCTGTGGGAGGAGGGACCCAACAGCT 321

     127 heGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuLys 143
         |||||||||||||||||:::||||||||||||||||||   |||||||||
     322 TCCAGCCCTCCGACCCGGTCCTCACCATGGCCGTGGACAGCGGGCTCAAG 371

     144 AspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrpLy 160
         ||||||||||||:::|||||||||||||||||||||||||||||||||::
     372 GATGACTTGGTGTTCGGGGACCCCAACGCGCCCCGGTTCGTGCTGTGGGA 421

     160 sGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePheA 177
         :|||||||||||||||||||||||||||      |||||||||||||||:
     422 GGGGAAGCTGAGGCCGGTGCCGTCGAAGCCAGGCGACCTGCCTTTCTTCA 471

     177 spLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleGly 193
         ::|||||||||||||||||||||||||||||||||:::|||   :::|||
     472 GCCTCATGAGTATCCCTGGGAAGCTCAGGGCCGGCCTTGGCGCGCTCGGC 521

     194 LeuArgProSerProProGlyHisGluGluSerValGluGlnPheValAr 210
         :::||||||   |||||||||:::|||||||||||||||:::||||||||
     522 ATTCGCCCACCTCCTCCAGGGCGCGAGGAGTCGGTGGAGGAGTTTGTGCG 571

     210 gArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCysS 227
         |||||||||||||:::||||||||||||||||||||||||||||||||||
     572 CCGCAACCTCGGTGCCGAGGTCTTTGAGCGCCTCATCGAGCCTTTCTGCT 621

     227 erGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaPhe 243
         |||||||||||:::||||||||||||||||||||||||||||||||||||
     622 CAGGTGTATATGCTGGTGATCCTTCGAAGCTTAGTATGAAGGCTGCATTT 671

     244 GlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGlyTh 260
         ||||||||||||:::|||||||||   |||||||||||||||||||||||
     672 GGGAAGGTCTGGAGGTTGGAGGAGATTGGAGGTAGTATTATTGGTGGAAC 721

     260 rPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAspP 277
         |:::|||||||||::::::::::::::::::||||||   ||||||||||
     722 CATCAAGGCGATTCAGGATAAAGGGAAGAACCCCAAACCGCCAAGGGATC 771

     277 roArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysGly 293
         |||||||||||   ||||||||||||||||||:::|||||||||||||||
     772 CCCGACTTCCGGCACCAAAGGGACAGACGGTGGCATCTTTCAGGAAGGGT 821

     294 LeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeuLy 310
         |||   |||||||||:::||||||::::::|||||||||||||||:::||
     822 CTAGCCATGCTCCCGAATGCCATCGCATCTAGGCTGGGTAGTAAAGTCAA 871

     310 sLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTyrH 327
         ||||||||||||||||:::||||||||||||:::::::::   ||||||
     872 GCTGTCATGGAAGCTTACGAGCATTACAAAGGCGGACAACCAAGGATATG 921

     327 isLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgSer 343
           |||   ||||||||||||||||||:::||||||:::|||::::::|||
     922 TATTAGGTTATGAAACACCAGAAGGACTTGTTTCAGTGCAGGCTAAAAGT 971

     344 IleValMetThrValProSerTyrValAlaSerAsnIleLeuArgProLe 360
         ::::::||||||:::||||||||||||||||||:::||||||||||||||
     972 GTTATCATGACCATCCCGTCATATGTTGCTAGTGATATCTTGCGCCCACT 1021

     360 uSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProValG 377
         ||||:::   ||||||||||||||||||:::||||||||||||||||||:
    1022 TTCAATTGATGCAGCAGATGCACTCTCAAAATTCTATTATCCGCCAGTTG 1071

     377 lyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgLeu 393
         :::|||||||||:::|||||||||:::||||||||||||   |||   |||
```

```
1072 CTGCTGTAACTGTTTCATATCCAAAAGAAGCTATTAGAAAAGAATGCTTA 1121

 394 ValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGl 410
     :::||||||||||||:::||||||||||||||||||||||||:::|||||
1122 ATTGATGGGGAGCTCCAGGGTTTCGGCCAGTTGCATCCACGTAGCCAAGG 1171

 410 yValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnArgA 427
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1172 AGTCGAGACTTTAGGGACAATATATAGCTCTTCTCTCTTTCCTAATCGTG 1221

 427 laProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysAsn 443
     |||||   |||||||||||||||||||||||||||||||||:::   |||
1222 CTCCTGCTGGAAGAGTGTTACTTCTGAACTATATCGGGGGTTCTACAAAT 1271

 444 ProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAspAr 460
           |||:::|||||||||||||||:::||||||   :::||||||||
1272 ACAGGGATCGTCTCCAAGACTGAGAGTGACTTAGTAGGAGCCGTTGACCG 1321

 460 gAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuValV 477
     ||||||||||||||||||||||:::|||:::|||   |||||||||::::
1322 TGACCTCAGAAAAATGTTGATAAACCCTAGAGCAGCAGACCCTTTAGCAT 1371

 477 alGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyHis 493
     ::|||||||||||||||||||||||||||||||||||||||:::||||||
1372 TAGGGGTTCGAGTGTGGCCACAAGCAATACCACAGTTTTTGATTGGGCAC 1421

 494 LeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGl 510
     ||||||   |||::::::||||||:::|||::::::::::::|||   ::
1422 CTTGATCGCCTTGCTGCTGCAAAATCTGCACTGGGCCAAGGCGGCTACGA 1471

 510 uGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyArgC 527
     :|||||||||||||||||||||||||||:::|||||||||||||||||||
1472 CGGGTTGTTCCTAGGAGGAAACTACGTCGCAGGAGTTGCCTTGGGCCGAT 1521

 527 ysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuSer 543
     ||:::|||||||||||||||   ||||||:::|||:::   ||||||:::
1522 GCATCGAGGGTGCGTACGAGAGTGCCTCACAAGTATCTGACTTCTTGACC 1571

 544 ArgTyrAlaTyrLys 548
     :::||||||||||||
1572 AAGTATGCCTACAAG 1586

seq_name: gb_pat2:AX084736

seq_documentation_block:
LOCUS       AX084736     1691 bp    DNA             PAT       09-MAR-2001
DEFINITION  Sequence 5 from Patent WO0112825.
ACCESSION   AX084736
VERSION     AX084736.1  GI:13274989
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
            Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1691)
  AUTHORS   Johnson,M.A., Volrath,S.L., Heifetz,P.B. and Law,M.D.
  TITLE     Herbicide-tolerant protoporphyrinogen oxidase
  JOURNAL   Patent: WO 0112825-A 5 22-FEB-2001;
            Syngenta Participations AG (CH)
FEATURES             Location/Qualifiers
     source          1. .1691
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
     CDS             <1. .1446
                     /note="Maize protox-1 c-DNA (not full-length)"
                     /codon_start=1
                     /protein_id="CAC34010.1"
                     /db_xref="GI:13274990"
                     /translation="ADCVVVGGGISGLCTAQALATRHGVGDVLVTEARARPGGNITTV
                     ERPEEGYLWEEGPNSFQPSDPVLTMAVDSGLKDDLVFGDPNAPRFVLWEGKLRPVPSK
                     PADLPFFDLMSIPGKLRAGLGALGIRPPPPGREESVEEFVRRNLGAEVFERLIEPFCS
                     GVYAGDPSKLSMKAAFGKVWRLEETGGSIIGGTIKTIQERSKNPKPPRDARLPKPKGQ
                     TVASFRKGLAMLPNAITSSLGSKVKLSWKLTSITKSDDKGYVLEYETPEGVVSVQAKS
                     VIMTIPSYVASNILRPLSSDAADALSRFYYPPVAAVTVSYPKEAIRKECLIDGELQGF
                     GQLHPRSQGVETLGTIYSSSLFPNRAPDGRVLLLNYIGGATNTGIVSKTESELVEAVD
                     RDLRKMLINSTAVDPLVLGVRVWPQAIPQFLVGHLDLLEAAKAALDRGGYDGLFLGGN
                     YVAGVALGRCVEGAYESASQISDFLTKYAYK"
BASE COUNT      414 a    393 c    452 g    432 t
ORIGIN


alignment_scores:
           Quality: 1978.00              Length:     480
             Ratio:  4.455                 Gaps:       2
 Percent Similarity:  92.500   Percent Identity: 78.750

alignment_block:
 US-09-508-418-2 x AX084736   ..

 Align seg 1/1  to: AX084736  from: 1  to: 1691

      73 AspCysValIleValGlyAlaGlyIleSerGlyLeuCysIleAlaGlnVa 89
         |||||||||:::||||||:::||||||||||||||||||   ||||||::
       4 GACTGCGTCGTGGTGGGCGGAGGCATCAGTGGCCTCTGCACCGCGCAGGC 53

      89 lMetSerAlaAsnTyr......ProAsnLeuMetValThrGluAlaArgA 104
         ::::::::::::::::         :::::::::||||||||||||||| 
      54 GCTGGCCACGCGGCACGGCGTCGGGGACGTGCTTGTCACGGAGGCCCGCG 103

     104 spArgAlaGlyGlyAsnIleThrThrValGluArg......AspGlyTyr 118
           |||   |||||||||||||||||||||||||||      :::||||||
     104 CCCGCCCCGGCGGCAACATTACCACCGTCGAGCGCCCCGAGGAAGGGTAC 153

     119 LeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetLeuTh 135
         ||||||||||||||||||||||||||||||||||||||||||:::|||||
     154 CTCTGGGAGGAGGGTCCCAACAGCTTCCAGCCCTCCGACCCCGTTCTCAC 203

     135 rMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyAspProA 152
         |||||||||||||   |||||||||||||||||||||:::||||||||||
     204 CATGGCCGTGGACAGCGGACTGAAGGATGACTTGGTTTTGGGGACCCAA 253

     152 snAlaProArgPheValLeuTrpLysGlyLysLeuArgProValProSer 168
         ||||||||||||||||||||||||||:::|||||||||||||||||||||
     254 ACGCGCCGCGTTTCGTGCTGTGGGAGGGGAAGCTGAGGCCCGTGCCATCC 303

     169 LysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyLysLe 185
         |||   :::|||||||||||||||||||||||||||||||||||||||||
     304 AAGCCCGCCGACCTCCCGTTCTTCGATCTCATGAGCATCCCAGGGAAGCT 353

     185 uArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGlyHisG 202
         ||||||||||:::|||   :::|||:::||||||   |||||||||:::|
     354 CAGGGCCGGTCTAGGCGCGCTTGGCATCCGCCCGCCTCCTCCAGGCCGCG 403

     202 luGluSerValGluGlnPheValArgArgAsnLeuGlyGlyGluValPhe 218
         ||||||||||||||:::|||||||||||||||||||||:::|||||||||
     404 AAGAGTCAGTGGAGGAGTTCGTGCGCCGCAACCTCGGTGCTGAGGTCTTT 453

     219 GluArgLeuIleGluProPheCysSerGlyValTyrValGlyAspProSe 235
         ||||||||||||||||||||||||||||||||||||:::|||||||||||
     454 GAGCGCCTCATTGAGCCTTTCTGCTCAGGTGTCTATGCTGGTGATCCTTC 503

     235 rLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLeuGluGluT 252
         |||||||||||||||||||||||||||||||||||||:::||||||||||
     504 TAAGCTCAGCATGAAGGCTGCATTTGGGAAGGTTTGGCGGTTGGAAGAAA 553

     252 hrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysGluArgSer 268
         ||||||||||||||||||||||||||:::|||:::|||:::|||||||||
     554 CTGGAGGTAGTATTATTGGTGGAACCATCAAGACAATTCAGGAGAGGAGC 603

     269 SerThrProLysAlaProArgAspProArgLeuProLysProLysGlyGl 285
         ::::::||||||   |||||||||   |||||||||||||||||||||||
     604 AAGAATCCAAAACCACCGAGGGATGCCCGCCTTCCGAAGCCAAAAGGGCA 653

     285 nThrValGlySerPheArgLysGlyLeuArgMetLeuProAspAlaIleS 302
         |||||||:::||||||||||||||||||   |||||||||:::||||||:
```

```
 654 GACAGTTGCATCTTTCAGGAAGGGTCTTGCCATGCTTCCAAATGCCATTA 703

 302 erAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSerSerIle 318
     :::::   ||||||||||||:::||||||||||||||||||:::||||||
 704 CATCCAGCTTGGGTAGTAAAGTCAAACTATCATGGAAACTCACGAGCATT 753

 319 ThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrProGluGl 335
     |||||||||:::      ||||||   |||   |||||||||||||||||
 754 ACAAAATCAGATGACAAGGGATATGTTTTGGAGTATGAAACGCCAGAAGG 803

 335 yValValSerLeuGlnSerArgSerIleValMetThrValProSerTyrV 352
     ||||||||||:::|||::::::|||::::::||||||:::||||||||||
 804 GGTTGTTTCGGTGCAGGCTAAAAGTGTTATCATGACTATTCCATCATATG 853

 352 alAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAlaLeu 368
     |||||||||||||||||||||||||||||      |||||||||||||||
 854 TTGCTAGCAACATTTTGCGTCCACTTTCAAGCGATGCTGCAGATGCTCTA 903

 369 SerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrProGl 385
     |||:::||||||||||||||||||:::|||||||||:::|||||||||::
 904 TCAAGATTCTATTATCCACCGGTTGCTGCTGTAACTGTTTCGTATCCAAA 953

 385 nGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGlyPheG 402
     :||||||||||||   |||   |||:::||||||||||||:::|||||||
 954 GGAAGCAATTAGAAAAGAATGCTTAATTGATGGGGAACTCCAGGGCTTTG 1003

 402 lyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrIleTyr 418
     |||||||||||||||||:::||||||||||||||||||||||||||||||
1004 GCCAGTTGCATCCACGTAGTCAAGGAGTTGAGACATTAGGAACAATATAC 1053

 419 SerSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuLeuLe 435
     ||||||||||||||||||||||||||||||   |||||||||||||||||
1054 AGTTCCTCACTCTTTCCAAATCGTGCTCCTGACGGTAGGGTGTTACTTCT 1103

 435 uAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThrGluS 452
     |||||||||||||||||||   |||      |||:::|||||||||||||
1104 AAACTACATAGGAGGTGCTACAAACACAGGAATTGTTTCCAAGACTGAAA 1153

 452 erGlnLeuValGluValValAspArgAspLeuArgLysMetLeuIleLys 468
     ||:::|||||||||:::||||||||||||||||||||||||||||||:::
1154 GTGAGCTGGTCGAAGCAGTTGACCGTGACCTCCGAAAAATGCTTATAAAT 1203

 469 ProLysAlaGlnAspProLeuValValGlyValArgValTrpProGlnAl 485
           |||   ||||||||||||:::|||||||||||||||||||||||
1204 TCTACAGCAGTGGACCCTTTAGTCCTTGGTGTTCGAGTTTGGCCACAAGC 1253

 485 aIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAlaLysA 502
     |||||||||||||||||||||||||||||||   |||::::::|||||||
1254 CATACCTCAGTTCCTGGTAGGACATCTTGATCTTCTGGAAGCCGCAAAAG 1303

 502 laAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAsnTyr 518
     |||||::::::   :::|||   :::||||||||||||||||||||||||
1304 CTGCCCTGGACCGAGGTGGCTACGATGGGCTGTTCCTAGGAGGGAACTAT 1353

 519 ValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluValAl 535
     |||:::|||||||||||||||||||||||||||||||||||||||   ||
1354 GTTGCAGGAGTTGCCCTGGGCAGATGCGTTGAGGGCGCGTATGAAAGTGC 1403

 535 aSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
     ||||:::::::::   ||||||::::::||||||||||||
1404 CTCGCAAATATCTGACTTCTTGACCAAGTATGCCTACAAG 1443

seq_name: gb_pat1:A78959

seq_documentation_block:
LOCUS       A78959       1698 bp    DNA             PAT       20-OCT-1999
DEFINITION  Sequence 13 from Patent WO9829554.
ACCESSION   A78959
VERSION     A78959.1  GI:6092087
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
            Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1698)
  AUTHORS   Ishige,F. and Sato,R.
  TITLE     METHODS OF CONFERRING PPO-INHIBITING HERBICIDE RESISTANCE TO PLANTS
            BY GENE MANIPULATION
  JOURNAL   Patent: WO 9829554-A 13 09-JUL-1998;
            ISHIGE FUMIHARU (JP); SATO RYO (JP)
FEATURES             Location/Qualifiers
     source          1. .1698
                     /organism="Zea mays"
                     /strain="B73 INBRED"
                     /db_xref="taxon:4577"
     CDS             <2. .1453
                     /codon_start=1
                     /product="PROTOPORPHYRINOGEN OXIDASE"
                     /protein_id="CAB59133.1"
                     /db_xref="GI:6092088"
                     /translation="NSADCVVVGGGISGLCTAQALATRHGVGDVLVTEARARPGGNIT
                     TVERPEEGYLWEEGPNSFQPSDPVLTMAVDSGLKDDLVFGDPNAPRFVLWEGKLRPVP
                     SKPADLPFFDLMSIPGKLRAGLGALGIRPPPPGREESVEEFVRRNLGAEVFERLIEPF
                     CSGVYAGDPSKLSMKAAFGKVWRLEETGGSIIGGTIKTIQERSKNPKPPRDARLPKPK
                     GQTVASFRKGLAMLPNAITSSLGSKVKLSWKLTSITKSDDKGYVLEYETPEGVVSVQA
                     KSVIMTIPSYVASNILRPLSSDAADALSRFYYPPVAAVTVSYPKEAIRKECLIDGELQ
                     GFGQLHPRSQGVETLGTIYSSSLFPNRAPDGRVLLLNYIGGATNTGIVSKTESELVEA
                     VDRDLRKMLINSTAVDPLVLGVRVWPQAIPQFLVGHLDLLEAAKAALDRGGYDGLFLG
                     GNYVAGVALGRCVEGAYESASQISDFLTKYAYK"
BASE COUNT      416 a    396 c    454 g    432 t
ORIGIN


alignment_scores:
.            Quality: 1978.00              Length:    480
               Ratio:  4.455                 Gaps:      2
 Percent Similarity:  92.500   Percent Identity: 78.750

alignment_block:
 US-09-508-418-2 x A78959   ..

 Align seg 1/1  to: A78959  from: 1  to: 1698

      73 AspCysValIleValGlyAlaGlyIleSerGlyLeuCysIleAlaGlnVa 89
         |||||||||:::||||||:::||||||||||||||||||   ||||||::
      11 GACTGCGTCGTGGTGGGCGGAGGCATCAGTGGCCTCTGCACCGCGCAGGC 60

      89 lMetSerAlaAsnTyr......ProAsnLeuMetValThrGluAlaArgA 104
         ::::::::::::::::         :::::::::|||||||||||||||
      61 GCTGGCCACGCGGCACGGCGTCGGGGACGTGCTTGTCACGGAGGCCCGCG 110

     104 spArgAlaGlyGlyAsnIleThrThrValGluArg......AspGlyTyr 118
           |||.  |||||||||||||||||||||||||||      :::||||||
     111 CCCGCCCCGGCGGCAACATTACCACCGTCGAGCGCCCCGAGGAAGGGTAC 160

     119 LeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetLeuTh 135
         ||||||||||||||||||||||||||||||||||||||||||:::|||||
     161 CTCTGGGAGGAGGGTCCCAACAGCTTCCAGCCCTCCGACCCCGTTCTCAC 210

     135 rMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyAspProA 152
         |||||||||||||   |||||||||||||||||||||:::||||||||||
     211 CATGGCCGTGGACAGCGGACTGAAGGATGACTTGGTTTTTGGGGACCCAA 260

     152 snAlaProArgPheValLeuTrpLysGlyLysLeuArgProValProSer 168
         |||||||||||||||||||||||:::||||||||||||||||||||||||
     261 ACGCGCCGCGTTTCGTGCTGTGGGAGGGGAAGCTGAGGCCCGTGCCATCC 310

     169 LysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyLysLe 185
         |||   :::|||||||||||||||||||||||||||||||||||||||||
     311 AAGCCCGCCGACCTCCCGTTCTTCGATCTCATGAGCATCCCAGGGAAGCT 360

     185 uArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGlyHisG 202
         ||||||||||:::|||   :::|||:::||||||   |||||||||:::|
     361 CAGGGCCGGTCTAGGCGCGCTTGGCATCCGCCCGCCTCCTCCAGGCCGCG 410
```

```
 202 luGluSerValGluGlnPheValArgArgAsnLeuGlyGlyGluValPhe 218
     ||||||||||||||:::|||||||||||||||||||||:::|||||||||
 411 AAGAGTCAGTGGAGGAGTTCGTGCGCCGCAACCTCGGTGCTGAGGTCTTT 460

 219 GluArgLeuIleGluProPheCysSerGlyValTyrValGlyAspProSe 235
     ||||||||||||||||||||||||||||||||||||:::|||||||||||
 461 GAGCGCCTCATTGAGCCTTTCTGCTCAGGTGTCTATGCTGGTGATCCTTC 510

 235 rLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLeuGluGluT 252
     |||||||||||||||||||||||||||||||||||||:::||||||||||
 511 TAAGCTCAGCATGAAGGCTGCATTTGGGAAGGTTTGGCGGTTGGAAGAAA 560

 252 hrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysGluArgSer 268
     ||||||||||||||||||||||||||:::|||:::|||:::|||||||||
 561 CTGGAGGTAGTATTATTGGTGGAACCATCAAGACAATTCAGGAGAGGAGC 610

 269 SerThrProLysAlaProArgAspProArgLeuProLysProLysGlyGl 285
     ::::::||||||   |||||||||   |||||||||||||||||||||||
 611 AAGAATCCAAAACCACCGAGGGATGCCCGCCTTCCGAAGCCAAAAGGGCA 660

 285 nThrValGlySerPheArgLysGlyLeuArgMetLeuProAspAlaIleS 302
     |||||||:::||||||||||||||||||   |||||||||:::||||||:
 661 GACAGTTGCATCTTTCAGGAAGGGTCTTGCCATGCTTCCAAATGCCATTA 710

 302 erAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSerSerIle 318
     :::::   ||||||||||||:::||||||||||||||||||:::||||||
 711 CATCCAGCTTGGGTAGTAAAGTCAAACTATCATGGAAACTCACGAGCATT 760

 319 ThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrProGluGl 335
     |||||||||:::      ||||||   |||   |||||||||||||||||
 761 ACAAAATCAGATGACAAGGGATATGTTTTGGAGTATGAAACGCCAGAAGG 810

 335 yValValSerLeuGlnSerArgSerIleValMetThrValProSerTyrV 352
     ||||||||||:::|||::::::|||::::::||||||:::||||||||||
 811 GGTTGTTTCGGTGCAGGCTAAAAGTGTTATCATGACTATTCCATCATATG 860

 352 alAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAlaLeu 368
     |||||||||||||||||||||||||||||      |||||||||||||||
 861 TTGCTAGCAACATTTTGCGTCCACTTTCAAGCGATGCTGCAGATGCTCTA 910

 369 SerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrProGl 385
     |||:::||||||||||||||||||:::|||||||||:::|||||||||::
 911 TCAAGATTCTATTATCCACCGGTTGCTGCTGTAACTGTTTCGTATCCAAA 960

 385 nGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGlyPheG 402
     :||||||||||||   |||   |||:::||||||||||||:::|||||||
 961 GGAAGCAATTAGAAAAGAATGCTTAATTGATGGGGAACTCCAGGGCTTTG 1010

 402 lyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrIleTyr 418
     |||||||||||||||||:::||||||||||||||||||||||||||||||
1011 GCCAGTTGCATCCACGTAGTCAAGGAGTTGAGACATTAGGAACAATATAC 1060

 419 SerSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuLeuLe 435
     ||||||||||||||||||||||||||||||   |||||||||||||||||
1061 AGTTCCTCACTCTTTCCAAATCGTGCTCCTGACGGTAGGGTGTTACTTCT 1110

 435 uAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThrGluS 452
     |||||||||||||||||||   |||      |||:::|||||||||||||
1111 AAACTACATAGGAGGTGCTACAAACACAGGAATTGTTTCCAAGACTGAAA 1160

 452 erGlnLeuValGluValValAspArgAspLeuArgLysMetLeuIleLys 468
     ||:::|||||||||:::||||||||||||||||||||||||||||||:::
1161 GTGAGCTGGTCGAAGCAGTTGACCGTGACCTCCGAAAAATGCTTATAAAT 1210

 469 ProLysAlaGlnAspProLeuValValGlyValArgValTrpProGlnAl 485
           |||   ||||||||||||:::|||||||||||||||||||||||
1211 TCTACAGCAGTGGACCCTTTAGTCCTTGGTGTTCGAGTTTGGCCACAAGC 1260

 485 aIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAlaLysA 502
     |||||||||||||||||||||||||||||||   |||::::::|||||||
1261 CATACCTCAGTTCCTGGTAGGACATCTTGATCTTCTGGAAGCCGCAAAAG 1310

 502 laAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAsnTyr 518
     |||||::::::   :::|||   :::||||||||||||||||||||||||
1311 CTGCCCTGGACCGAGGTGGCTACGATGGGCTGTTCCTAGGAGGGAACTAT 1360

 519 ValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluValAl 535
     |||:::|||||||||||||||||||||||||||||||||||||||   ||
1361 GTTGCAGGAGTTGCCCTGGGCAGATGCGTTGAGGGCGCGTATGAAAGTGC 1410

 535 aSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
     ||||:::::::::   ||||||::::::||||||||||||
1411 CTCGCAAATATCTGACTTCTTGACCAAGTATGCCTACAAG 1450
```

```
OM of: US-09-508-418-2 to: N_Geneseq_0601:*   out_format : pfs

Date: Jul 3, 2001 11:40 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09508418/runat_03072001_100017_24474/app_query.fasta_1.613
 -DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US09508418_@CGN1_1_222 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-508-418-2
Query length: 548
Database: N_Geneseq_0601:*
Database sequences: 730101
Database length: 313950809
Search time (sec): 112.980000



score_list:
Sequence              Strd Orig    ZScore   EScore Len  ! Documentation  ..
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX28279 + 2836.00 4513.22  2.0e-243  1874  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX28278 + 2832.00 4506.82  4.6e-243  1874  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04313 + 2113.00 3356.55  5.4e-179  1826  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86123 + 2113.00 3356.55  5.4e-179  1826  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76579 + 2113.00 3356.55  5.4e-179  1826  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT:AAT11673 + 2077.00 3299.59  8.1e-176  1719  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04304 + 2077.00 3299.59  8.1e-176  1719  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86129 + 2077.00 3299.59  8.1e-176  1719  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV07251 + 2077.00 3299.59  8.1e-176  1719  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC47180 + 2075.00 3296.57  1.2e-175  1690  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76571 + 2068.00 3285.19  5.1e-175  1719  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ45271 + 2062.50 3276.95  1.5e-174  1632  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04309 + 2062.50 3275.61  1.8e-174  1847  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86122 + 2062.50 3275.61  1.8e-174  1847  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76576 + 2062.50 3275.61  1.8e-174  1847  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04315 + 2006.50 3186.37  1.6e-169  1784  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86125 + 2006.50 3186.37  1.6e-169  1784  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76581 + 2006.50 3186.37  1.6e-169  1784  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX02093 + 1988.50 3156.83  7.3e-168  1911  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04308 + 1987.00 3155.01  9.2e-168  1811  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86121 + 1987.00 3155.01  9.2e-168  1811  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76575 + 1987.00 3155.01  9.2e-168  1811  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04306 + 1978.00 3141.35  5.3e-167  1691  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86128 + 1978.00 3141.35  5.3e-167  1691  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV07253 + 1978.00 3141.35  5.3e-167  1691  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76573 + 1978.00 3141.35  5.3e-167  1691  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT:AAT11675 + 1978.00 3141.30  5.3e-167  1698  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86136 + 1952.50 3099.96  1.1e-164  1784  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04314 + 1933.50 3068.82  5.8e-163  1910  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86124 + 1933.50 3068.82  5.8e-163  1910  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76580 + 1933.50 3068.82  5.8e-163  1910  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04317 + 1823.00 2893.98  3.2e-153  1590  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86127 + 1823.00 2893.98  3.2e-153  1590  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76583 + 1823.00 2893.98  3.2e-153  1590  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ32119 + 1530.50 2424.35  4.6e-127  1838  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ45282 + 1530.50 2424.35  4.6e-127  1838  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX02095 + 1298.00 2033.50  2.7e-105  10455
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04316 + 1261.00 1997.48  2.7e-103  1224  !
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86126 + 1261.00 1997.48  2.7e-103  1224  !
/SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76582 + 1261.00 1997.48  2.7e-103  1224  !
```

```
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX02103 +  524.00  832.86  2.0e-38   313
/SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX01529 +  522.50  811.89  3.0e-37   1745
/SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:AAT95392 +  490.50  761.21  2.0e-34   1662
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04305 +  490.50  760.72  2.1e-34   1738
/SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86130 +  490.50  760.72  2.1e-34   1738

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX28279

seq_documentation_block:
ID   AAX28279 standard; cDNA to mRNA; 1874 BP.
XX
AC   AAX28279;
XX
DT   16-JUN-1999  (first entry)
XX
DE   N. tabacum protoporphyrinogen oxidase mutant coding sequence.
XX
KW   Protoporphyrinogen oxidase; light-requiring herbicide tolerance; enzyme;
KW   ds.
XX
OS   Nicotiana tabacum.
XX
PN   WO9913087-A1.
XX
PD   18-MAR-1999.
XX
PF   10-SEP-1998;   98WO-JP04064.
XX
PR   11-SEP-1997;   97JP-0265084.
XX
PA   (NIPY ) NIPPON NOYAKU CO LTD.
XX
PI   Hirooka T,  Horikoshi M,  Mametsuka K;
XX
DR   WPI; 1999-215067/18.
DR   P-PSDB; AAY05201.
XX
PT   Protoporphyrinogen oxidase tolerant to light-requiring herbicides,
PT   useful for constructing plants with such properties
XX
PS   Example 5; Page 44-49; 56pp; Japanese.
XX
CC   This sequence encodes the Nicotiana tabacum protoporphyrinogen oxidase
CC   of the invention, which is tolerant to light-requiring herbicides. The
CC   enzyme and its derivatives can be applied to construct plants with high
CC   tolerance to light-requiring herbicides, useful in agriculture.
XX
SQ   Sequence 1874 BP; 509 A; 382 C; 461 G; 522 T; 0 other;


alignment_scores:
            Quality: 2836.00             Length:    548
              Ratio:  5.185                Gaps:      0
 Percent Similarity:  99.818   Percent Identity: 99.818

alignment_block:
 US-09-508-418-2 x AAX28279   ..

 Align seg 1/1  to: AAX28279  from: 1  to: 1874

       1 MetThrThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSe 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      26 ATGACAACAACTCCCATCGCCAATCATCCTAATATTTTCACTCACCAGTC 75

      17 rSerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheS 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      76 GTCGTCATCGCCATTGGCATTCTTAAACCGTACGAGTTTCATCCCTTTCT 125

      34 erSerIleSerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArg 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     126 CTTCAATCTCCAAGCGCAATAGTGTCAATTGCAATGGCTGGAGAACACGA 175

      51 CysSerValAlaLysAspTyrThrValProSerSerAlaValAspGlyGl 67
```

```
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   176 TGCTCCGTTGCCAAAGATTACACAGTTCCTTCCTCAGCGGTCGACGGCGG 225

    67 yProAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyL 84
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   226 ACCCGCCGCGGAGCTGGACTGTGTTATAGTTGGAGCAGGAATTAGTGGCC 275

    84 euCysIleAlaGlnValMetSerAlaAsnTyrProAsnLeuMetValThr 100
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   276 TCTGCATTGCGCAGGTGATGTCCGCTAATTACCCCAATTTGATGGTAACC 325

   101 GluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgAspGl 117
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   326 GAGGCGAGAGATCGTGCCGGTGGCAACATAACGACTGTGGAAAGAGACGG 375

   117 yTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetL 134
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   376 CTATTTGTGGGAAGAAGGTCCCAACAGTTTCCAGCCGTCCGATCCTATGT 425

   134 euThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyAsp 150
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   426 TGACTATGGCAGTAGATTGTGGATTGAAGGATGATTTGGTGTTGGGAGAT 475

   151 ProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProValPr 167
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   476 CCTAATGCGCCCCGTTTCGTTTTGTGGAAGGGTAAATTAAGGCCCGTCCC 525

   167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyL 184
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   526 CTCAAAACTCACTGATCTTCCCTTTTTGATTTGATGAGCATTCCTGGCA 575

   184 ysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGly 200
       ||||||||||||||||||||   |||||||||||||||||||||||||||
   576 AGTTGAGAGCTGGTTTTGGTGCCATTGGCCTCCGCCCTTCACCTCCAGGT 625

   201 HisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyGluVa 217
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   626 CATGAGGAATCAGTTGAGCAGTTCGTGCGTCGTAATCTTGGTGGCGAAGT 675

   217 lPheGluArgLeuIleGluProPheCysSerGlyValTyrValGlyAspP 234
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   676 CTTTGAACGCTTGATAGAACCATTTTGTTCTGGTGTTTATGTTGGTGATC 725

   234 roSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLeuGlu 250
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   726 CCTCAAAACTGAGTATGAAAGCAGCATTTGGGAAAGTTTGGAAGTTGGAA 775

   251 GluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysGluAr 267
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   776 GAAACTGGTGGTAGCATTATTGGAGGAACCTTTAAAGCAATAAAGGAGAG 825

   267 gSerSerThrProLysAlaProArgAspProArgLeuProLysProLysG 284
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   826 ATCCAGTACACCTAAAGCGCCCCGCGATCCGCGTTTACCTAAACCAAAAG 875

   284 lyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAspAla 300
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   876 GACAGACAGTTGGATCATTCAGGAAGGGTCTCAGAATGCTGCCGGATGCA 925

   301 IleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSerSe 317
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   926 ATCAGTGCAAGATTGGGAAGCAAATTAAAACTATCATGGAAGCTTTCTAG 975

   317 rIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrProG 334
       ||||||||||||||||||||||||||||||||||||||||||||||||||
   976 CATTACTAAGTCAGAAAAAGGAGGATATCACTTGACATACGAGACACCAG 1025

   334 luGlyValValSerLeuGlnSerArgSerIleValMetThrValProSer 350
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1026 AAGGAGTAGTTTCTCTTCAAAGTCGAAGCATTGTCATGACTGTGCCATCC 1075

   351 TyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAl 367
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1076 TATGTAGCAAGCAACATATTACGTCCTCTTTCGGTTGCCGCAGCAGATGC 1125

   367 aLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrP 384
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1126 ACTTTCAAATTTCTACTATCCCCCAGTTGGAGCAGTCACAATTTCATATC 1175

   384 roGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGly 400
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1176 CTCAAGAAGCTATTCGTGATGAGCGTCTGGTTGATGGTGAACTAAAGGGA 1225

   401 PheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrIl 417
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1226 TTTGGGCAGTTGCATCCACGTACACAGGGAGTGGAAACACTAGGAACGAT 1275

   417 eTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuL 434
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1276 ATATAGTTCATCACTCTTCCCTAACCGTGCCCCAAAAGGTCGGGTGCTAC 1325

   434 euLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThr 450
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1326 TCTTGAACTACATTGGAGGAGCAAAAAATCCTGAAATTTTGTCTAAGACG 1375

   451 GluSerGlnLeuValGluValValAspArgAspLeuArgLysMetLeuIl 467
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1376 GAGAGCCAACTTGTGGAAGTAGTTGATCGTGACCTCAGAAAAATGCTTAT 1425

   467 eLysProLysAlaGlnAspProLeuValValGlyValArgValTrpProG 484
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1426 AAAACCCAAAGCTCAAGATCCTCTTGTTGTGGGTGTGCGAGTATGGCCAC 1475

   484 lnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAla 500
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1476 AAGCTATCCCACAGTTTTTGGTTGGTCATCTGGATACGCTAAGTACTGCA 1525

   501 LysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAs 517
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1526 AAAGCTGCTATGAATGATAATGGGCTTGAAGGGCTGTTTCTTGGGGGTAA 1575

   517 nTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluV 534
       ||||||||||||||||||||||||||||||||||||||||||||||||||
  1576 TTATGTGTCAGGTGTAGCATTGGGGAGGTGTGTTGAAGGTGCTTATGAAG 1625

   534 alAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
       ||||||||||||||||||||||||||||||||||||||||||||
  1626 TTGCATCCGAGGTAACAGGATTTCTGTCTCGGTATGCATACAAA 1669

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX28278

seq_documentation_block:
ID   AAX28278 standard; cDNA to mRNA; 1874 BP.
XX
AC   AAX28278;
XX
DT   16-JUN-1999  (first entry)
XX
DE   N. tabacum protoporphyrinogen oxidase coding sequence.
XX
KW   Protoporphyrinogen oxidase; light-requiring herbicide tolerance; enzyme;
KW   ds.
XX
OS   Nicotiana tabacum.
XX
PN   WO9913087-A1.
XX
PD   18-MAR-1999.
XX
PF   10-SEP-1998;   98WO-JP04064.
XX
PR   11-SEP-1997;   97JP-0265084.
XX
PA   (NIPY ) NIPPON NOYAKU CO LTD.
XX
PI   Hirooka T,  Horikoshi M,  Mametsuka K;
```

```
XX
DR   WPI; 1999-215067/18.
DR   P-PSDB; AAY05200.
XX
PT   Protoporphyrinogen oxidase tolerant to light-requiring herbicides,
PT   useful for constructing plants with such properties
XX
PS   Example 5; Page 34-39; 56pp; Japanese.
XX
CC   This sequence encodes the Nicotiana tabacum protoporphyrinogen oxidase
CC   of the invention, which is tolerant to light-requiring herbicides. The
CC   enzyme and its derivatives can be applied to construct plants with high
CC   tolerance to light-requiring herbicides, useful in agriculture.
XX
SQ   Sequence 1874 BP; 509 A; 383 C; 461 G; 521 T; 0 other;


alignment_scores:
           Quality: 2832.00             Length:    548
             Ratio:  5.177                Gaps:      0
 Percent Similarity:  99.818   Percent Identity: 99.635

alignment_block:
 US-09-508-418-2 x AAX28278   ..

 Align seg 1/1  to: AAX28278  from: 1  to: 1874

      1 MetThrThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSe 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     26 ATGACAACAACTCCCATCGCCAATCATCCTAATATTTCACTCACCAGTC 75

     17 rSerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheS 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     76 GTCGTCATCGCCATTGGCATTCTTAAACCGTACGAGTTTCATCCCTTTCT 125

     34 erSerIleSerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArg 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    126 CTTCAATCTCCAAGCGCAATAGTGTCAATTGCAATGGCTGGAGAACACGA 175

     51 CysSerValAlaLysAspTyrThrValProSerSerAlaValAspGlyGl 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    176 TGCTCCGTTGCCAAAGATTACACAGTTCCTTCCTCAGCGGTCGACGGCGG 225

     67 yProAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyL 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    226 ACCCGCCGCGGAGCTGGACTGTGTTATAGTTGGAGCAGGAATTAGTGGCC 275

     84 euCysIleAlaGlnValMetSerAlaAsnTyrProAsnLeuMetValThr 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    276 TCTGCATTGCGCAGGTGATGTCCGCTAATTACCCCAATTTGATGGTAACC 325

    101 GluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgAspGl 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    326 GAGGCGAGAGATCGTGCCGGTGGCAACATAACGACTGTGGAAAGAGACGG 375

    117 yTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetL 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    376 CTATTTGTGGGAAGAAGGTCCCAACAGTTTCCAGCCGTCCGATCCTATGT 425

    134 euThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyAsp 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    426 TGACTATGGCAGTAGATTGTGGATTGAAGGATGATTTGGTGTTGGGAGAT 475

    151 ProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProValPr 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    476 CCTAATGCGCCCCGTTTCGTTTTGTGGAAGGGTAAATTAAGGCCCGTCCC 525

    167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyL 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    526 CTCAAAACTCACTGATCTTCCCTTTTTTGATTTGATGAGCATTCCTGGCA 575

    184 ysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGly 200
        ||||||||||||||||||||   |||||||||||||||||||||||||||
    576 AGTTGAGAGCTGGTTTTGGTGCCATTGGCCTCCGCCCTTCACCTCCAGGT 625

    201 HisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyGluVa 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    626 CATGAGGAATCAGTTGAGCAGTTCGTGCGTCGTAATCTTGGTGGCGAAGT 675

    217 lPheGluArgLeuIleGluProPheCysSerGlyValTyrValGlyAspP 234
        ||||||||||||||||||||||||||||||||||||||||:::|||||||
    676 CTTTGAACGCTTGATAGAACCATTTTGTTCTGGTGTTTATGCTGGTGATC 725

    234 roSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLeuGlu 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    726 CCTCAAAACTGAGTATGAAAGCAGCATTTGGGAAAGTTTGGAAGTTGGAA 775

    251 GluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysGluAr 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    776 GAAACTGGTGGTAGCATTATTGGAGGAACCTTTAAAGCAATAAAGGAGAG 825

    267 gSerSerThrProLysAlaProArgAspProArgLeuProLysProLysG 284
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    826 ATCCAGTACACCTAAAGCGCCCCGCGATCCGCGTTTACCTAAACCAAAAG 875

    284 lyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAspAla 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    876 GACAGACAGTTGGATCATTCAGGAAGGGTCTCAGAATGCTGCCGGATGCA 925

    301 IleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSerSe 317
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    926 ATCAGTGCAAGATTGGGAAGCAAATTAAAACTATCATGGAAGCTTTCTAG 975

    317 rIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrProG 334
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    976 CATTACTAAGTCAGAAAAAGGAGGATATCACTTGACATACGAGACACCAG 1025

    334 luGlyValValSerLeuGlnSerArgSerIleValMetThrValProSer 350
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1026 AAGGAGTAGTTTCTCTTCAAAGTCGAAGCATTGTCATGACTGTGCCATCC 1075

    351 TyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAl 367
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1076 TATGTAGCAAGCAACATATTACGTCCTCTTTCGGTTGCCGCAGCAGATGC 1125

    367 aLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrP 384
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1126 ACTTTCAAATTTCTACTATCCCCCAGTTGGAGCAGTCACAATTTCATATC 1175

    384 roGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGly 400
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1176 CTCAAGAAGCTATTCGTGATGAGCGTCTGGTTGATGGTGAACTAAAGGGA 1225

    401 PheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrIl 417
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1226 TTTGGGCAGTTGCATCCACGTACACAGGGAGTGGAAACACTAGGAACGAT 1275

    417 eTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuL 434
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1276 ATATAGTTCATCACTCTTCCCTAACCGTGCCCCAAAAGGTCGGGTGCTAC 1325

    434 euLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThr 450
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1326 TCTTGAACTACATTGGAGGAGCAAAAAATCCTGAAATTTTGTCTAAGACG 1375

    451 GluSerGlnLeuValGluValValAspArgAspLeuArgLysMetLeuIl 467
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1376 GAGAGCCAACTTGTGGAAGTAGTTGATCGTGACCTCAGAAAAATGCTTAT 1425

    467 eLysProLysAlaGlnAspProLeuValValGlyValArgValTrpProG 484
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1426 AAAACCCAAAGCTCAAGATCCTCTTGTTGTGGGTGTGCGAGTATGGCCAC 1475

    484 lnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAla 500
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1476 AAGCTATCCCACAGTTTTTGGTTGGTCATCTGGATACGCTAAGTACTGCA 1525
```

```
     501 LysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAs 517
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1526 AAAGCTGCTATGAATGATAATGGGCTTGAAGGGCTGTTTCTTGGGGGTAA 1575

     517 nTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluV 534
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1576 TTATGTGTCAGGTGTAGCATTGGGGAGGTGTGTTGAAGGTGCTTATGAAG 1625

     534 alAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
         ||||||||||||||||||||||||||||||||||||||||||||
    1626 TTGCATCCGAGGTAACAGGATTTCTGTCTCGGTATGCATACAAA 1669

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04313

seq_documentation_block:
ID   AAV04313 standard; cDNA; 1826 BP.
XX
AC   AAV04313;
XX
DT   20-APR-1998  (first entry)
XX
DE   Cotton protox-1 cDNA.
XX
KW   Protoporphyrinogen oxidase-1; protox-1; promoter; cotton;
KW   herbicide resistance; breeding programme; probe; gene isolation;
KW   genomic mapping; ss.
XX
OS   Gossypium hirsutum.
XX
FH   Key             Location/Qualifiers
FT   CDS             31..1650
FT                   /*tag=  a
FT                   /product= protox-1
XX
PN   WO9732028-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US03343.
XX
PR   21-JUN-1996;   96US-0020003.
PR   28-FEB-1996;   96US-0012705.
PR   28-FEB-1996;   96US-0013612.
XX
PA   (NOVS ) NOVARTIS AG.
XX
PI   Johnson MA,  Volrath SL,  Ward ER;
XX
DR   WPI; 1997-489209/45.
DR   P-PSDB; AAW41609.
XX
PT   DNA containing a plant proto-porphyrinogen oxidase gene promoter -
PT   optionally linked to a heterologous gene, especially to express
PT   herbicide-resistant enzymes, and plants containing such constructs
XX
PS   Disclosure; Pages 78-80; 114pp; English.
XX
CC   The present sequence encodes cotton protoporphyrinogen oxidase-1
CC   (protox-1).
CC   The protox-1 promoter can be used to express herbicide resistant
CC   enzymes, specifically protox, i.e. a plant tissue, plant or progeny
CC   containing a chimeric gene of the promoter and a heterologous
CC   coding sequence. The plant can also be used in breeding programmes.
CC   Also hybridising fragments of the protox coding sequence can be
CC   used as probes, e.g. to isolate related genes or for genomic
CC   mapping.
XX
SQ   Sequence 1826 BP; 471 A; 399 C; 439 G; 517 T; 0 other;


alignment_scores:
           Quality: 2113.00            Length:    506
```

```
              Ratio:  4.477              Gaps:     1
 Percent Similarity:  93.281   Percent Identity: 79.249

alignment_block:
 US-09-508-418-2 x AAV04313   ..

 Align seg 1/1  to: AAV04313  from: 1  to: 1826

      47 TrpArgThrArgCysSerValAlaLysAspTyrThrValProSerSerAl 63
         ::::::   |||||||||:::|||:::      |||:::   ||||||
     130 TTCAAGCTCCGATGCTCCCTCGCCGAGGGTCCCACGATTTCCTCATCTAA 179

      63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
          :::|||||||||   ::::::      ||||||||||||||||||:::|
     180 AATCGACGGGGGAGAATCATCCATCGCGGATTGCGTCATCGTTGGAGGTG 229

      80 lyIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr...... 94
         ||||||||||||||||||||||||||::::::::::::::::::
     230 GTATCAGTGGACTTTGCATTGCTCAAGCTCTCGCCACCAAGCACCGTGAC 279

      95 ......ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAs 109
                  |||::::::|||||||||||||||||||||:::||||||||
     280 GTCGCTTCCAATGTGATTGTGACGGAGGCCAGAGACCGTGTTGGTGGCAA 329

     109 nIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnS 126
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     330 CATCACTACCGTTGAGAGAGATGGATATCTGTGGGAAGAAGGCCCCAACA 379

     126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
         ||||||||||||||||||||:::||||||||||||||||||   ||||||
     380 GTTTTCAGCCCTCCGATCCTATTCTAACCATGGCCGTGGATAGTGGATTG 429

     143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     430 AAGGACGATTTGGTTTTAGGTGACCCTAATGCACCGCGATTTGTACTATG 479

     159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
         |:::|||||||||||||||||||||||||||   ||||||||||||||||
     480 GGAGGGAAAACTAAGGCCTGTGCCCTCCAAGCCAACCGACTTGCCGTTTT 529

     176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
         |||||||||||||||||   ||||||||||||||||||||||||   |||
     530 TTGATTTGATGAGCATTGCTGGAAAACTTAGGGCTGGGTTCGGGGCTATT 579

     193 GlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVa 209
         |||:::||||||   |||||||||:::|||||||||||||||:::|||||
     580 GGCATTCGGCCTCCCCCTCCGGGTTATGAAGAATCGGTGGAGGAGTTTGT 629

     209 lArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC 226
         ||||||||||||||||:::|||||||||||||||:::|||||||||||||
     630 GCGCCGTAATCTTGGTGCTGAGGTTTTTGAACGCTTTATTGAACCATTTT 679

     226 ysSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
         ||||||||||||||:::|||||||||||||||||||||||||||||||||
     680 GTTCAGGTGTTTATGCAGGGGATCCTTCAAAATTAAGCATGAAAGCAGCA 729

     243 PheGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGl 259
         ||||||:::||||||||||||||||||   ||||||||||||||||||||
     730 TTTGGAAGAGTATGGAAGCTAGAAGAGATTGGTGGCAGCATCATTGGTGG 779

     259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
         ||||||||||:::|||:::||||||::::::|||||||||   |||||||
     780 CACTTTCAAGACAATCCAGGAGAGAAATAAGACACCTAAGCCACCCAGAG 829

     276 spProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 292
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     830 ACCCGCGTCTGCCAAAACCGAAGGGCCAAACAGTTGGATCTTTTAGGAAG 879

     293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
         ||||||   |||||||||:::||||||:::      |||||||||:::::
     880 GGACTTACCATGCTGCCTGAGGCAATTGCTAACAGTTTGGGTAGCAATGT 929

     309 uLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyT 326
```

```
      :|||||||||||||||||||||||||||||||||      :::|||||||
 930 AAAATTATCTTGGAAGCTTTCCAGTATTACCAAATTGGGCAATGGAGGGT 979

 326 yrHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
     ||:::||||||:::|||||||||||||||:::||||||||||||||||||
 980 ATAACTTGACATTTGAAACACCTGAAGGAATGGTATCTCTTCAGAGTAGA 1029

 343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
     |||:::|||||||||:::||||||:::||||||||||||:::|||:::||
1030 AGTGTTGTAATGACCATTCCATCCCATGTTGCCAGTAACTTGTTGCATCC 1079

 359 oLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProV 376
     |||||||:::|||||||||||||||||||||:::||||||||||||||||
1080 TCTCTCGGCTGCTGCTGCAGATGCATTATCCCAATTTTATTATCCTCCAG 1129

 376 alGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArg 392
     ||::::::||||||:::|||||||||:::||||||||||||   |||
1130 TTGCATCAGTCACAGTCTCCTATCCAAAAGAAGCCATTCGAAAAGAATGT 1179

 393 LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGl 409
     |||:::|||||||||||||||||||||||||||||||||||||||:::||
1180 TTGATTGATGGTGAACTTAAGGGGTTTGGCCAGTTGCACCCACGCAGCCA 1229

 409 nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnA 426
     ||||:::|||||||||||||||||||||||||||||||||||||||||||
1230 AGGAATTGAAACTTTAGGGACGATATACAGTTCATCACTTTTCCCCAATC 1279

 426 rgAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
     ||||||||:::||||||||||||||||||||||||||||||||||||
1280 GAGCTCCATCTGGCAGGGTGTTGCTCTTGAACTACATAGGAGGAGCTACC 1329

 443 AsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAs 459
     |||      ||||||||||||||||||::::::|||||||||:::|||||
1330 AACACTGGAATTTTGTCCAAGACTGAAGGGGAACTTGTAGAAGCAGTTGA 1379

 459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
     |||||||||||||||||||||||||:::|||:::|||:::||||||||||
1380 TCGTGATTTGAGAAAAATGCTTATAAATCCTAATGCAAAGGATCCTCTTG 1429

 476 alValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGly 492
     ||:::||||||||||||||||||:::||||||||||||||||||||||||
1430 TTTTGGGTGTAAGAGTATGGCCAAAAGCCATTCCACAGTTCTTGGTTGGT 1479

 493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe 509
     |||||||||   |||::::::||||||   |||::::::|||:::|||::
1480 CATTTGGATCTCCTTGATAGTGCAAAAATGGCTCTCAGGGATTCTGGGTT 1529

 509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
     ::::||||||||||||||||||||||||||||||||||||||||||||||
1530 TCATGGACTGTTTCTTGGGGGCAACTATGTATCTGGTGTGGCATTAGGAC 1579

 526 rgCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 542
     |||||||||||||||||||||||||||||:::||||||      ||||||
1580 GGTGTGTGGAAGGTGCTTACGAGGTTGCAGCTGAAGTGAAGGAATTCCTG 1629

 543 SerArgTyrAlaTyrLys 548
     |||:::||||||||||||
1630 TCACAATATGCATACAAA 1647
```

```
seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86123

seq_documentation_block:
ID   AAT86123 standard; cDNA; 1826 BP.
XX
AC   AAT86123;
XX
DT   01-MAR-1998  (first entry)
XX
DE   Cotton protoporphyrinogen oxidase (protox-1) cDNA clone pWDC-15.
XX
KW   Protox-1; protoporphyrinogen oxidase; inhibitor; cotton;
KW   herbicide tolerance; herbicide resistance; transgenic plant; ss.
XX
OS   Gossypium hirsutum.
XX
FH   Key             Location/Qualifiers
FT   CDS             31..1650
FT                   /*tag=  a
XX
PN   WO9732011-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US03313.
XX
PR   21-JUN-1996;   96US-0020003.
PR   28-FEB-1996;   96US-0012705.
PR   28-FEB-1996;   96US-0013612.
XX
PA   (NOVS ) NOVARTIS AG.
XX
PI   Heifetz PB,  Johnson MA,  Potter SL,  Volrath SL,  Ward ER;
XX
DR   WPI; 1997-448683/41.
DR   P-PSDB; AAW25740.
XX
PT   New DNA encoding plant protoporphyrinogen oxidase enzyme  - and
PT   herbicide resistant mutants, useful to prepare plants resistant to
PT   herbicide which therefore kills undesired vegetation only
XX
PS   Claim 8; Page 137-139; 196pp; English.
XX
CC   This cDNA clone codes for cotton protoporphyrinogen oxidase
CC   (Protox-1, see AAW25740).  It was isolated from a dark-grown
CC   cotyledon library using maize Protox-1 cDNA (see AAT86128) as probe.
CC   Cotton Protox-1 in pBluescript SK vector has been deposited as
CC   pWDC-15 (NRRL AAB21594).  The isolated cDNA can be mutated so that
CC   it encodes a modified protox enzyme that is resistant to protox
CC   inhibitors and hence is herbicide tolerant.  Plants, especially
CC   crop plants, may be engineered for resistance to protox inhibitors
CC   via mutation of the native protox gene to a resistant form, or they
CC   may be transformed with a gene encoding an inhibitor-resistant form
CC   of a plant protox enzyme, including claimed forms from wheat,
CC   soybean, cotton, sugarbeet, oilseed rape, rice and sorghum (see
CC   AAW25738-45).  Application of herbicide therefore kills undesired
CC   vegetation only.  Plant protox nucleic acids can also be used as
CC   probes and PCR primers, as selectable markers in plant cell
CC   transformation methods, and for recombinant production of protox
CC   enzymes in host cells.
XX
SQ   Sequence 1826 BP; 471 A; 399 C; 439 G; 517 T; 0 other;


alignment_scores:
           Quality: 2113.00              Length:    506
             Ratio:  4.477                 Gaps:      1
 Percent Similarity:  93.281   Percent Identity: 79.249

alignment_block:
 US-09-508-418-2 x AAT86123    ..

 Align seg 1/1  to: AAT86123  from: 1  to: 1826

      47 TrpArgThrArgCysSerValAlaLysAspTyrThrValProSerSerAl 63
         ::::::   |||||||||:::|||:::      |||:::   ||||||
     130 TTCAAGCTCCGATGCTCCCTCGCCGAGGGTCCCACGATTTCCTCATCTAA 179

      63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
          :::|||||||||   ::::::      ||||||||||||||||||:::|
     180 AATCGACGGGGGAGAATCATCCATCGCGGATTGCGTCATCGTTGGAGGTG 229

      80 lyIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr...... 94
         ||||||||||||||||||||||||||::::::::::::::::::
     230 GTATCAGTGGACTTTGCATTGCTCAAGCTCTCGCCACCAAGCACCGTGAC 279

      95 ......ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAs 109
```

```
              |||::::::|||||||||||||||||||||:::||||||||
 280 GTCGCTTCCAATGTGATTGTGACGGAGGCCAGAGACCGTGTTGGTGGCAA 329

 109 nIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnS 126
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 330 CATCACTACCGTTGAGAGAGATGGATATCTGTGGGAAGAAGGCCCCAACA 379

 126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
     ||||||||||||||||||||:::||||||||||||||||||   ||||||
 380 GTTTTCAGCCCTCCGATCCTATTCTAACCATGGCCGTGGATAGTGGATTG 429

 143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 430 AAGGACGATTTGGTTTTAGGTGACCCTAATGCACCGCGATTTGTACTATG 479

 159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
     |:::|||||||||||||||||||||||||||   ||||||||||||||||
 480 GGAGGGAAAACTAAGGCCTGTGCCCTCCAAGCCAACCGACTTGCCGTTTT 529

 176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
     |||||||||||||||||   ||||||||||||||||||||||||   |||
 530 TTGATTTGATGAGCATTGCTGGAAAACTTAGGGCTGGGTTCGGGGCTATT 579

 193 GlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVa 209
     |||:::||||||   |||||||||:::|||||||||||||||:::|||||
 580 GGCATTCGGCCTCCCCCTCCGGGTTATGAAGAATCGGTGGAGGAGTTTGT 629

 209 lArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC 226
     ||||||||||||||||:::|||||||||||||||:::|||||||||||||
 630 GCGCCGTAATCTTGGTGCTGAGGTTTTTGAACGCTTTATTGAACCATTTT 679

 226 ysSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
     ||||||||||||||:::|||||||||||||||||||||||||||||||||
 680 GTTCAGGTGTTTATGCAGGGGATCCTTCAAAATTAAGCATGAAAGCAGCA 729

 243 PheGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGl 259
     ||||||:::||||||||||||||||||   ||||||||||||||||||||
 730 TTTGGAAGAGTATGGAAGCTAGAAGAGATTGGTGGCAGCATCATTGGTGG 779

 259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
     ||||||||||:::|||:::||||||::::::|||||||||   |||||||
 780 CACTTTCAAGACAATCCAGGAGAGAAATAAGACACCTAAGCCACCCAGAG 829

 276 spProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 292
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 830 ACCCGCGTCTGCCAAAACCGAAGGGCCAAACAGTTGGATCTTTTAGGAAG 879

 293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
     ||||||   |||||||||:::||||||:::      |||||||||:::::
 880 GGACTTACCATGCTGCCTGAGGCAATTGCTAACAGTTTGGGTAGCAATGT 929

 309 uLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyT 326
     :|||||||||||||||||||||||||||||||||      :::|||||||
 930 AAAATTATCTTGGAAGCTTTCCAGTATTACCAAATTGGGCAATGGAGGGT 979

 326 yrHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
     ||:::||||||:::|||||||||||||||:::||||||||||||||||||
 980 ATAACTTGACATTTGAAACACCTGAAGGAATGGTATCTCTTCAGAGTAGA 1029

 343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
     |||:::|||||||||:::||||||:::||||||||||||:::|||:::||
1030 AGTGTTGTAATGACCATTCCATCCCATGTTGCCAGTAACTTGTTGCATCC 1079

 359 oLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProV 376
     |||||||:::|||||||||||||||||||||:::||||||||||||||||
1080 TCTCTCGGCTGCTGCTGCAGATGCATTATCCCAATTTTATTATCCTCCAG 1129

 376 alGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArg 392
     ||::::::||||||:::|||||||||:::||||||||||||   |||
1130 TTGCATCAGTCACAGTCTCCTATCCAAAAGAAGCCATTCGAAAAGAATGT 1179

 393 LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGl 409
     |||:::|||||||||||||||||||||||||||||||||||||||:::||
1180 TTGATTGATGGTGAACTTAAGGGGTTTGGCCAGTTGCACCCACGCAGCCA 1229

 409 nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnA 426
     ||||:::|||||||||||||||||||||||||||||||||||||||||||
1230 AGGAATTGAAACTTTAGGGACGATATACAGTTCATCACTTTTCCCCAATC 1279

 426 rgAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
     ||||||||:::||||||||||||||||||||||||||||||||||||
1280 GAGCTCCATCTGGCAGGGTGTTGCTCTTGAACTACATAGGAGGAGCTACC 1329

 443 AsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAs 459
     |||      ||||||||||||||||||::::::|||||||||:::|||||
1330 AACACTGGAATTTTGTCCAAGACTGAAGGGGAACTTGTAGAAGCAGTTGA 1379

 459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
     |||||||||||||||||||||||||:::|||:::|||:::||||||||||
1380 TCGTGATTTGAGAAAAATGCTTATAAATCCTAATGCAAAGGATCCTCTTG 1429

 476 alValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGly 492
     ||:::||||||||||||||||||:::||||||||||||||||||||||||
1430 TTTTGGGTGTAAGAGTATGGCCAAAAGCCATTCCACAGTTCTTGGTTGGT 1479

 493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe 509
     |||||||||   |||::::::||||||   |||::::::|||:::|||::
1480 CATTTGGATCTCCTTGATAGTGCAAAAATGGCTCTCAGGGATTCTGGGTT 1529

 509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
     ::::||||||||||||||||||||||||||||||||||||||||||||||
1530 TCATGGACTGTTTCTTGGGGGCAACTATGTATCTGGTGTGGCATTAGGAC 1579

 526 rgCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 542
     |||||||||||||||||||||||||||||:::||||||      ||||||
1580 GGTGTGTGGAAGGTGCTTACGAGGTTGCAGCTGAAGTGAAGGAATTCCTG 1629

 543 SerArgTyrAlaTyrLys 548
     |||:::||||||||||||
1630 TCACAATATGCATACAAA 1647

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76579

seq_documentation_block:
ID   AAF76579 standard; cDNA; 1826 BP.
XX
AC   AAF76579;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Cotton protoporphyrinogen oxidase coding sequence SEQ ID NO: 15.
XX
KW   Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
KW   soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein; ss.
XX
OS   Gossypium hirsutum.
XX
PN   WO200112825-A1.
XX
PD   22-FEB-2001.
XX
PF   30-JUN-2000; 2000WO-EP06127.
XX
PR   13-AUG-1999;   99US-0373691.
XX
PA   (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX
PI   Johnson MA,  Volrath SL,  Heifetz PB,  Law MD;
XX
DR   WPI; 2001-234914/24.
DR   P-PSDB; AAB72907.
XX
PT   Plant DNA molecules encoding herbicide-tolerant forms of
PT   protoporphyrinogen oxidase which are useful for rationally designing
PT   new inhibitory herbicides and for producing herbicide-tolerant
PT   transgenic plants and seeds -
```

```
XX
PS   Claim 2; Page 180-184; 228pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
CC   sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
CC   oxidase (protox) enzyme. Examples of these mutants are shown in
CC   AAB72920-AAB72926. They are useful as they enable the production of
CC   herbicide-tolerant plants and seeds. The present sequence is a protox
CC   coding sequence.
XX
SQ   Sequence 1826 BP; 471 A; 399 C; 439 G; 517 T; 0 other;


alignment_scores:
          Quality: 2113.00          Length:    506
            Ratio:  4.477             Gaps:      1
 Percent Similarity:  93.281   Percent Identity: 79.249

alignment_block:
 US-09-508-418-2 x AAF76579   ..

 Align seg 1/1  to: AAF76579  from: 1  to: 1826

     47 TrpArgThrArgCysSerValAlaLysAspTyrThrValProSerSerAl 63
        ::::::   |||||||||:::|||:::      |||:::   ||||||
    130 TTCAAGCTCCGATGCTCCCTCGCCGAGGGTCCCACGATTTCCTCATCTAA 179

     63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
         :::|||||||||   ::::::      ||||||||||||||||||:::|
    180 AATCGACGGGGGAGAATCATCCATCGCGGATTGCGTCATCGTTGGAGGTG 229

     80 lyIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr...... 94
        ||||||||||||||||||||||||||::::::::::::::::::
    230 GTATCAGTGGACTTTGCATTGCTCAAGCTCTCGCCACCAAGCACCGTGAC 279

     95 ......ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAs 109
                 |||::::::|||||||||||||||||||||:::||||||||
    280 GTCGCTTCCAATGTGATTGTGACGGAGGCCAGAGACCGTGTTGGTGGCAA 329

    109 nIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnS 126
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    330 CATCACTACCGTTGAGAGAGATGGATATCTGTGGGAAGAAGGCCCCAACA 379

    126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
        ||||||||||||||||||||:::||||||||||||||||||   ||||||
    380 GTTTTCAGCCCTCCGATCCTATTCTAACCATGGCCGTGGATAGTGGATTG 429

    143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    430 AAGGACGATTTGGTTTTAGGTGACCCTAATGCACCGCGATTTGTACTATG 479

    159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
        |:::|||||||||||||||||||||||||||   ||||||||||||||||
    480 GGAGGGAAAACTAAGGCCTGTGCCCTCCAAGCCAACCGACTTGCCGTTTT 529

    176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
        |||||||||||||||||   ||||||||||||||||||||||||   |||
    530 TTGATTTGATGAGCATTGCTGGAAAACTTAGGGCTGGGTTCGGGGCTATT 579

    193 GlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVa 209
        |||:::||||||   |||||||||:::|||||||||||||||:::|||||
    580 GGCATTCGGCCTCCCCCTCCGGGTTATGAAGAATCGGTGGAGGAGTTTGT 629

    209 lArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC 226
        ||||||||||||||||:::|||||||||||||||:::|||||||||||||
    630 GCGCCGTAATCTTGGTGCTGAGGTTTTTGAACGCTTTATTGAACCATTTT 679

    226 ysSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
        ||||||||||||||:::|||||||||||||||||||||||||||||||||
    680 GTTCAGGTGTTTATGCAGGGGATCCTTCAAAATTAAGCATGAAAGCAGCA 729

    243 PheGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGl 259
        ||||||:::||||||||||||||||||   ||||||||||||||||||||
    730 TTTGGAAGAGTATGGAAGCTAGAAGAGATTGGTGGCAGCATCATTGGTGG 779

    259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
        ||||||||||:::|||:::||||||::::::|||||||||   |||||||
    780 CACTTTCAAGACAATCCAGGAGAGAAATAAGACACCTAAGCCACCCAGAG 829

    276 spProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 292
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    830 ACCCGCGTCTGCCAAAACCGAAGGGCCAAACAGTTGGATCTTTTAGGAAG 879

    293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
        ||||||   |||||||||:::||||||:::      |||||||||:::::
    880 GGACTTACCATGCTGCCTGAGGCAATTGCTAACAGTTTGGGTAGCAATGT 929

    309 uLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyT 326
        :|||||||||||||||||||||||||||||||||      :::|||||||
    930 AAAATTATCTTGGAAGCTTTCCAGTATTACCAAATTGGGCAATGGAGGGT 979

    326 yrHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
        ||:::||||||:::|||||||||||||||:::||||||||||||||||||
    980 ATAACTTGACATTTGAAACACCTGAAGGAATGGTATCTCTTCAGAGTAGA 1029

    343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
        |||:::|||||||||:::||||||:::||||||||||||:::|||:::||
   1030 AGTGTTGTAATGACCATTCCATCCCATGTTGCCAGTAACTTGTTGCATCC 1079

    359 oLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProV 376
        |||||||:::|||||||||||||||||||||:::||||||||||||||||
   1080 TCTCTCGGCTGCTGCTGCAGATGCATTATCCCAATTTTATTATCCTCCAG 1129

    376 alGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArg 392
        ||::::::||||||:::|||||||||:::||||||||||||   |||
   1130 TTGCATCAGTCACAGTCTCCTATCCAAAAGAAGCCATTCGAAAAGAATGT 1179

    393 LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGl 409
        |||:::|||||||||||||||||||||||||||||||||||||||:::||
   1180 TTGATTGATGGTGAACTTAAGGGGTTTGGCCAGTTGCACCCACGCAGCCA 1229

    409 nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnA 426
        ||||:::|||||||||||||||||||||||||||||||||||||||||||
   1230 AGGAATTGAAACTTTAGGGACGATATACAGTTCATCACTTTTCCCCAATC 1279

    426 rgAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
        ||||||||:::||||||||||||||||||||||||||||||||||||
   1280 GAGCTCCATCTGGCAGGGTGTTGCTCTTGAACTACATAGGAGGAGCTACC 1329

    443 AsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAs 459
        |||      ||||||||||||||||||::::::|||||||||:::|||||
   1330 AACACTGGAATTTTGTCCAAGACTGAAGGGGAACTTGTAGAAGCAGTTGA 1379

    459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
        |||||||||||||||||||||||||:::|||:::|||:::||||||||||
   1380 TCGTGATTTGAGAAAAATGCTTATAAATCCTAATGCAAAGGATCCTCTTG 1429

    476 alValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGly 492
        ||:::||||||||||||||||||:::||||||||||||||||||||||||
   1430 TTTTGGGTGTAAGAGTATGGCCAAAAGCCATTCCACAGTTCTTGGTTGGT 1479

    493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe 509
        |||||||||   |||::::::||||||   |||::::::|||:::|||::
   1480 CATTTGGATCTCCTTGATAGTGCAAAAATGGCTCTCAGGGATTCTGGGTT 1529

    509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
        ::::||||||||||||||||||||||||||||||||||||||||||||||
   1530 TCATGGACTGTTTCTTGGGGGCAACTATGTATCTGGTGTGGCATTAGGAC 1579

    526 rgCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 542
        |||||||||||||||||||||||||||||:::||||||      ||||||
   1580 GGTGTGTGGAAGGTGCTTACGAGGTTGCAGCTGAAGTGAAGGAATTCCTG 1629

    543 SerArgTyrAlaTyrLys 548
        |||:::||||||||||||
```

```
   1630 TCACAATATGCATACAAA 1647

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT:AAT11673

seq_documentation_block:
ID   AAT11673 standard; cDNA; 1719 BP.
XX
AC   AAT11673;
XX
DT   02-JUL-1996  (first entry)
XX
DE   DNA encoding a protein having protoporphyrinogen oxidase activity.
XX
KW   Protoporphyrinogen oxidase; PPO; herbicide; resistance; mutant
KW   variegate porphyria; ss.
XX
OS   Arabidopsis sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             31..1644
FT                   /*tag= a
FT                   /product= Protoporphyrinogen oxidase.
XX
PN   WO9534659-A1.
XX
PD   21-DEC-1995.
XX
PF   08-JUN-1995;   95WO-IB00452.
XX
PR   16-JUN-1994;   94US-0261198.
XX
PA   (CIBA ) CIBA GEIGY AG.
XX
PI   Volrath S,  Ward ER;
XX
DR   WPI; 1996-049687/05.
DR   P-PSDB; AAR90295.
XX
PT   DNA encoding eukaryotic proto:porphyrinogen oxidase and herbicide
PT   resistant mutants  - used to make herbicide resistant plants and for
PT   diagnosis and treatment of variegate porphyria
XX
PS   Claim 1; Page 69-71; 118pp; English.
XX
CC   DNA encoding a modified protoporphyrinogen oxidase (PPO) or a
CC   chimeric gene comprising a promoter (pref. active in a plant) linked
CC   to such a DNA or DNA encoding a wild type PPO can be used (1) to
CC   impart herbicide resistance to plants; (2) for treating and
CC   diagnosing deficient PPO activity in animals (esp. variegate
CC   porphyria); and (3) for the production of recombinant PPO which is
CC   useful as an assay reagent and in rational design of new inhibitory
CC   herbicides.  Herbicide resistant PPO genes can also be used to
CC   select plants transformed with a transgene and probes derived from
CC   the genes can be used to quantify levels of PPO mRNA.
XX
SQ   Sequence 1719 BP; 460 A; 351 C; 433 G; 475 T; 0 other;


alignment_scores:
           Quality: 2077.00            Length:    539
             Ratio:  4.230               Gaps:      4
 Percent Similarity:  91.095   Percent Identity: 75.139

alignment_block:
 US-09-508-418-2 x AAT11673   ..

 Align seg 1/1  to: AAT11673  from: 1  to: 1719

      19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
         |||      |||::::::|||:::   |||:::         :::   ||
      28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

      35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
    |:::||||||   |||   ::::::   |||   :::         |||
 78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

 49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
      |||||||||||||||         ||||||   ||||||   ::::::
125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

 66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
    ||||||   ::::::   :::      ||||||||||||||||||:::||
175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

 80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
    |||||||||||||||||||||||||::::::::::::::::::
225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274

 95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
         |||||||||:::||||||||||||:::||||||:::|||||||||
275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324

110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
    |||   |||   |||::::::|||:::|||||||||||||||||||||||
325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
    |||||||||||||||||||||||||||||||:::||||||   |||||||
375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
    ||||||||||||||||||||||||||:::|||||||||||||||||||||
425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
    :::|||||||||||||||||||||||||||||||||||||||||||||||
475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524

176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
    ||||||||||||||||   ||||||:::|||||||||||||||   :::|
525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
    ||:::||||||||||||||||||:::|||||||||||||||:::||||||
575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
    |||||||||||||||   ||||||||||||||||||||||||||||||||
625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674

226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
    |||||||||||||:::||||||||||||||||||||||||||||||||||
675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
    |||||||||||||||||||||||::::::|||||||||||||||||||||
725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
    |||||||||||||||:::||||||:::::::::|||||||||   |||||
775 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
    |||||||||||||||||||:::||||||||||||||||||||||||||||
825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 874

293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
    |||||||||||||||||:::|||||||||||||||||||||||||||:::
875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
    |||||||||||||||||||||:::|||||||||   |||:::||||||||
925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
    |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
```

```
  975 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1024

  343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
      ||:::||||||||||||||||||:::|||||||||::::::|||||||||
 1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

  360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
      ||||||   :::||||||:::|||||||||::::::||||||||||||||
 1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

  376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
      |:::||||||:::||||||||||||:::||||||||||||   |||   |
 1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

  393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
      ||:::|||||||||||||||||||||||||||||||||||||||||||||
 1175 TGATAGATGGTGAACTAAAGGGTTTGGGCAATTGCATCCACGCACGCAA 1224

  410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274

  426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
      |||||||   ||||||:::||||||||||||||||||||||||:::   |
 1275 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

  443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
      ||      ||||||||||||:::|||::::::|||||||||:::||||||
 1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

  460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
      ||||||||||||||||||||||||||||||::::::   |||||||||
 1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

  476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
       :::||||||||||||||||||||||||||||||||||||||||||||||
 1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

  493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
      ||:::|||   |||:::|||||||||::::::::::::::::::|||
 1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

  510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
      ||||||||||||||||||||||||||||||:::|||||||||||||||||
 1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

  526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
      ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
 1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

  543 erArgTyrAlaTyrLys 548
      |||||||||||||||||
 1625 CACGGTACGCTTACAAG 1641

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04304

seq_documentation_block:
ID   AAV04304 standard; cDNA; 1719 BP.
XX
AC   AAV04304;
XX
DT   20-APR-1998  (first entry)
XX
DE   Arabidopsis thaliana protox-1 cDNA.
XX
KW   Protoporphyrinogen oxidase-1; protox-1; promoter;
KW   herbicide resistance; breeding programme; probe; gene isolation;
KW   genomic mapping; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             31..1644
```

```
FT                   /*tag= a
FT                   /product= protox-1
XX
PN   WO9732028-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US03343.
XX
PR   21-JUN-1996;   96US-0020003.
PR   28-FEB-1996;   96US-0012705.
PR   28-FEB-1996;   96US-0013612.
XX
PA   (NOVS ) NOVARTIS AG.
XX
PI   Johnson MA,  Volrath SL,  Ward ER;
XX
DR   WPI; 1997-489209/45.
DR   P-PSDB; AAW41603.
XX
PT   DNA containing a plant proto-porphyrinogen oxidase gene promoter -
PT   optionally linked to a heterologous gene, especially to express
PT   herbicide-resistant enzymes, and plants containing such constructs
XX
PS   Example 1; Pages 36-39; 114pp; English.
XX
CC   The present sequence encodes Arabidopsis thaliana
CC   protoporphyrinogen oxidase-1 (protox-1).
CC   The protox-1 promoter can be used to express herbicide resistant
CC   enzymes, specifically protox, i.e. a plant tissue, plant or progeny
CC   containing a chimeric gene of the promoter and a heterologous
CC   coding sequence. The plant can also be used in breeding programmes.
CC   Also hybridising fragments of the protox coding sequence can be
CC   used as probes, e.g. to isolate related genes or for genomic
CC   mapping.
XX
SQ   Sequence 1719 BP; 460 A; 351 C; 433 G; 475 T; 0 other;


alignment_scores:
            Quality: 2077.00            Length:    539
              Ratio:  4.230               Gaps:      4
 Percent Similarity:  91.095   Percent Identity: 75.139

alignment_block:
 US-09-508-418-2 x AAV04304    ..

 Align seg 1/1  to: AAV04304  from: 1  to: 1719

   19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
      |||      |||::::::|||:::   |||:::         :::   ||
   28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

   35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
      |:::||||||   |||   ::::::   |||   :::         |||
   78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

   49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
        |||||||||||||||         ||||||   ||||||   ::::::
  125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

   66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
      ||||||   ::::::   :::      ||||||||||||||||||:::||
  175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

   80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
      |||||||||||||||||||||||||::::::::::::::::::
  225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274

   95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
           |||||||||:::||||||||||||:::||||||:::|||||||||
  275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324
```

```
 110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
     |||   |||   |||::::::|||:::|||||||||||||||||||||||
 325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

 126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
     |||||||||||||||||||||||||||||||:::||||||   |||||||
 375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

 143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
     ||||||||||||||||||||||||||:::|||||||||||||||||||||
 425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

 160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
     :::|||||||||||||||||||||||||||||||||||||||||||||||
 475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524

 176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
     ||||||||||||||||   ||||||:::|||||||||||||||   :::|
 525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

 193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
     ||:::||||||||||||||||||:::|||||||||||||||:::||||||
 575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

 210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
     |||||||||||||||   ||||||||||||||||||||||||||||||||
 625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674

 226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
     |||||||||||||:::||||||||||||||||||||||||||||||||||
 675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

 243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
     |||||||||||||||||||||||::::::|||||||||||||||||||||
 725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

 260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
     |||||||||||||||:::||||||:::::::::|||||||||   |||||
 775 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

 276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
     |||||||||||||||||||:::||||||||||||||||||||||||||||
 825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 874

 293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
     |||||||||||||||||:::|||||||||||||||||||||||||||:::
 875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

 310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
     |||||||||||||||||||||:::|||||||||   |||:::||||||||
 925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

 326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
     |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
 975 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1024

 343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
     ||:::||||||||||||||||||:::|||||||||::::::|||||||||
1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
     ||||||   :::||||||:::|||||||||::::::||||||||||||||
1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

 376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
     |:::||||||:::||||||||||||:::||||||||||||   |||   |
1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

 393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
     ||:::|||||||||||||||||||||||||||||||||||||||||||||
1175 TGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAA 1224

 410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274

 426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
     |||||||   ||||||:::||||||||||||||||||||||||:::   |
1275 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

 443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
     ||      ||||||||||||:::|||::::::|||||||||:::||||||
1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

 460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
     ||||||||||||||||||||||||||||||::::::   |||||||||
1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

 476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
      :::||||||||||||||||||||||||||||||||||||||||||||||
1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

 493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
     ||:::|||   |||:::|||||||||::::::::::::::::::|||
1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

 510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
     ||||||||||||||||||||||||||||||:::|||||||||||||||||
1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

 526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
     ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

 543 erArgTyrAlaTyrLys 548
     |||||||||||||||||
1625 CACGGTACGCTTACAAG 1641
```

```
seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86129

seq_documentation_block:
ID   AAT86129 standard; cDNA; 1719 BP.
XX
AC   AAT86129;
XX
DT   01-MAR-1998  (first entry)
XX
DE   Arabidopsis protoporphyrinogen oxidase (protox-1) cDNA pWDC-2.
XX
KW   Protox-1; protoporphyrinogen oxidase; inhibitor;
KW   herbicide tolerance; herbicide resistance; transgenic plant; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             31..1644
FT                   /*tag=  a
XX
PN   WO9732011-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US03313.
XX
PR   21-JUN-1996;   96US-0020003.
PR   28-FEB-1996;   96US-0012705.
PR   28-FEB-1996;   96US-0013612.
XX
PA   (NOVS ) NOVARTIS AG.
XX
PI   Heifetz PB,  Johnson MA,  Potter SL,  Volrath SL,  Ward ER;
XX
DR   WPI; 1997-448683/41.
DR   P-PSDB; AAW25746.
XX
PT   New DNA encoding plant protoporphyrinogen oxidase enzyme  - and
```

```
PT   herbicide resistant mutants, useful to prepare plants resistant to
PT   herbicide which therefore kills undesired vegetation only
XX
PS   Example 2; Page 95-99; 196pp; English.
XX
CC   This cDNA clone codes for Arabidopsis protoporphyrinogen oxidase
CC   (Protox-1, see AAW25746).  The clone has been deposited in pWDC-2 as
CC   NRRL B-21238.  The isolated cDNA can be mutated so that it encodes
CC   a modified protox enzyme that is resistant to protox inhibitors and
CC   hence is herbicide tolerant.  Plants, especially crop plants, may be
CC   engineered for resistance to protox inhibitors via mutation of the
CC   native protox gene to a resistant form, or they may be transformed
CC   with a gene encoding an inhibitor-resistant form of a plant protox
CC   enzyme, including claimed forms from wheat, soybean, cotton,
CC   sugarbeet, oilseed rape, rice, sorghum and Arabidopsis (see
CC   AAW25738-48).  Application of herbicide therefore kills undesired
CC   vegetation only.  Plant protox nucleic acids can also be used as
CC   probes and PCR primers, as selectable markers in plant cell
CC   transformation methods, and for recombinant production of protox
CC   enzymes in host cells.  Arabidopsis Protox-1 cDNA has been used in
CC   the identification of soybean (see AAT86122), sugarbeet (see AAT86124)
CC   and rape (see AAT86125) Protox-1 cDNA.
XX
SQ   Sequence 1719 BP; 460 A; 351 C; 433 G; 475 T; 0 other;

alignment_scores:
           Quality: 2077.00            Length:    539
             Ratio:  4.230               Gaps:      4
 Percent Similarity: 91.095   Percent Identity: 75.139

alignment_block:
 US-09-508-418-2 x AAT86129   ..

 Align seg 1/1  to: AAT86129  from: 1  to: 1719

     19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
        |||      |||::::::|||:::   |||:::         :::   ||
     28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

     35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
        |:::||||||   |||   ::::::   |||   :::         |||
     78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

     49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
          |||||||||||||||         ||||||   ||||||   ::::::
    125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

     66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
        ||||||   ::::::   :::      ||||||||||||||||||:::||
    175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

     80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
        |||||||||||||||||||||||||::::::::::::::::::
    225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274

     95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
             |||||||||:::||||||||||||:::||||||:::|||||||||
    275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324

    110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
        |||   |||   |||::::::|||:::|||||||||||||||||||||||
    325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

    126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
        |||||||||||||||||||||||||||||||:::||||||   |||||||
    375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

    143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
        ||||||||||||||||||||||||||:::|||||||||||||||||||||
    425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

    160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
        :::|||||||||||||||||||||||||||||||||||||||||||||||
    475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524

    176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
        ||||||||||||||||   ||||||:::|||||||||||||||   :::|
    525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

    193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
        ||:::||||||||||||||||||:::|||||||||||||||:::||||||
    575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

    210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
        |||||||||||||||   ||||||||||||||||||||||||||||||||
    625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674

    226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
        |||||||||||||:::||||||||||||||||||||||||||||||||||
    675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

    243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
        |||||||||||||||||||||||::::::|||||||||||||||||||||
    725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

    260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
        |||||||||||||||:::||||||:::::::::|||||||||   |||||
    775 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

    276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
        |||||||||||||||||||:::||||||||||||||||||||||||||||
    825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 874

    293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
        |||||||||||||||||:::|||||||||||||||||||||||||||:::
    875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

    310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
        |||||||||||||||||||||:::|||||||||   |||:::||||||||
    925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

    326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
        |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
    975 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1024

    343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
        ||:::||||||||||||||||||:::|||||||||::::::|||||||||
   1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

    360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
        ||||||   :::||||||:::|||||||||::::::||||||||||||||
   1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

    376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
        |:::||||||:::||||||||||||:::||||||||||||   |||   |
   1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

    393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
        ||:::|||||||||||||||||||||||||||||||||||||||||||||
   1175 TGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAA 1224

    410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274

    426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
        |||||||   ||||||:::||||||||||||||||||||||||:::   |
   1275 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

    443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
        ||      ||||||||||||:::|||::::::|||||||||:::||||||
   1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

    460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
        ||||||||||||||||||||||||||||||::::::   |||||||||
```

```
    1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

     476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
          :::||||||||||||||||||||||||||||||||||||||||||||||
    1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

     493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
         ||:::|||   |||:::|||||||||::::::::::::::::::|||
    1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

     510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
         ||||||||||||||||||||||||||||||:::|||||||||||||||||
    1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

     526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
         ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
    1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

     543 erArgTyrAlaTyrLys 548
         |||||||||||||||||
    1625 CACGGTACGCTTACAAG 1641

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV07251

seq_documentation_block:
ID   AAV07251 standard; cDNA; 1719 BP.
XX
AC   AAV07251;
XX
DT   14-AUG-1998  (first entry)
XX
DE   Arabidopsis cDNA encoding proto-porphyrinogen oxidase-1 (protox-1).
XX
KW   Proto-porphyrinogen oxidase; protox; herbicide tolerant; resistance;
KW   inhibitor; variegate porphyria; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             31..1644
FT                   /*tag= a
FT                   /product= Arabidopsis_protox-1
XX
PN   US5767373-A.
XX
PD   16-JUN-1998.
XX
PF   16-JUN-1994;   94US-0261198.
XX
PR   06-JUN-1995;   95US-0472028.
PR   16-JUN-1994;   94US-0261198.
XX
PA   (NOVS ) NOVARTIS FINANCE CORP.
XX
PI   Volrath S,  Ward ER;
XX
DR   WPI; 1998-361821/31.
DR   P-PSDB; AAW51347.
XX
PT   DNA encoding eukaryote herbicide resistant proto-porphyrinogen
PT   oxidase - useful for producing recombinant plants having functional
PT   enzyme, to be grown in the presence of herbicides
XX
PS   Claim 6; Columns 41-46; 43pp; English.
XX
CC   The invention relates to eukaryotic DNA sequences coding for native
CC   proto-porphyrinogen oxidase (protox) or modified forms of the enzyme
CC   which are herbicide tolerant. Plants having altered protox activity
CC   which confers tolerance to herbicides are also provided. These plants
CC   may be bred or engineered for resistance to protox inhibitors via
CC   mutation of the native protox gene to a resistant form or through
CC   increased levels of expression of the native protox gene, or they may be
CC   transformed with modified eukaryotic or prokaryotic protox coding
```

```
CC   sequences or wild type prokaryotic protox sequences which are herbicide
CC   tolerant. In the human condition variegate porphyria, an autosomal
CC   dominant disorder characterised by neuropsychiatric symptoms, decreased
CC   protox activity is detected. Thus the DNA molecules can be used to
CC   provide probes to detect and quantify protox levels in the diagnosis of
CC   diseases associated with the enzyme. The present sequence represents
CC   Arabidopsis cDNA encoding proto-porphyrinogen oxidase-1 (protox-1).
XX
SQ   Sequence 1719 BP; 460 A; 351 C; 433 G; 475 T; 0 other;


alignment_scores:
           Quality: 2077.00            Length:    539
             Ratio:  4.230               Gaps:      4
 Percent Similarity:  91.095   Percent Identity: 75.139

alignment_block:
 US-09-508-418-2 x AAV07251   ..

 Align seg 1/1  to: AAV07251  from: 1  to: 1719

      19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
         |||      |||::::::|||:::   |||:::         :::   ||
      28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

      35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
         |:::||||||   |||   ::::::   |||   :::         |||
      78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

      49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
           |||||||||||||||         ||||||   ||||||   ::::::
     125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

      66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
         ||||||   ::::::   :::      ||||||||||||||||||:::||
     175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

      80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
         |||||||||||||||||||||||||::::::::::::::::::
     225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274

      95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
              |||||||||:::||||||||||||:::||||||:::|||||||||
     275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324

     110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
         |||   |||   |||::::::|||:::|||||||||||||||||||||||
     325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

     126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
         |||||||||||||||||||||||||||||||:::||||||   |||||||
     375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

     143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
         ||||||||||||||||||||||||||:::|||||||||||||||||||||
     425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

     160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
         :::|||||||||||||||||||||||||||||||||||||||||||||||
     475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524

     176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
         ||||||||||||||||   ||||||:::|||||||||||||||   :::|
     525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

     193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
         ||:::||||||||||||||||||:::|||||||||||||||:::||||||
     575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

     210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
         |||||||||||||||   ||||||||||||||||||||||||||||||||
     625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674
```

```
 226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
     |||||||||||||:::||||||||||||||||||||||||||||||||||
 675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

 243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
     |||||||||||||||||||||||::::::|||||||||||||||||||||
 725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

 260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
     |||||||||||||||:::||||||:::::::::|||||||||   |||||
 775 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

 276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
     |||||||||||||||||||:::||||||||||||||||||||||||||||
 825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 874

 293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
     |||||||||||||||||:::|||||||||||||||||||||||||||:::
 875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

 310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
     |||||||||||||||||||||:::|||||||||   |||:::||||||||
 925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

 326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
     |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
 975 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1024

 343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
     ||:::||||||||||||||||||:::|||||||||::::::|||||||||
1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
     ||||||   :::||||||:::|||||||||::::::||||||||||||||
1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

 376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
     |:::||||||:::||||||||||||:::||||||||||||   |||   |
1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

 393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
     ||:::|||||||||||||||||||||||||||||||||||||||||||||
1175 TGATAGATGGTGAACTAAAGGGTTTGGGCAATTGCATCCACGCACGCAA 1224

 410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274

 426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
     |||||||   ||||||:::||||||||||||||||||||||||:::   |
1275 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

 443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
     ||      ||||||||||||:::|||::::::|||||||||:::||||||
1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

 460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
     ||||||||||||||||||||||||||||||::::::   |||||||||
1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

 476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
      :::||||||||||||||||||||||||||||||||||||||||||||||
1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

 493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
     ||:::|||   |||:::|||||||||::::::::::::::::::|||
1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

 510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
     ||||||||||||||||||||||||||||||:::|||||||||||||||||
1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

 526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
```

```
     ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

 543 erArgTyrAlaTyrLys 548
     |||||||||||||||||
1625 CACGGTACGCTTACAAG 1641

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC47180

seq_documentation_block:
ID   AAC47180 standard; DNA; 1690 BP.
XX
AC   AAC47180;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 52854.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
```

```
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
```

```
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

alignment_scores:
           Quality: 2075.00            Length:    536
             Ratio:  4.235               Gaps:      4
 Percent Similarity:  91.418   Percent Identity: 75.373

alignment_block:
 US-09-508-418-2 x AAC47180   ..
```

```
Align seg 1/1  to: AAC47180  from: 1  to: 1690

  22 LeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerIleSerLy 38
     |||::::::|||:::   |||:::         :::   |||:::|||||
   7 TTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTCGTTTTCGAA 56

  38 sArgAsnSerValAsnCysAsnGlyTrp.........ArgThrArgCysS 52
     |   |||   ::::::   |||   :::         |||   |||||||
  57 GCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGACTCCGTTGTT 103

  52 erValAlaLysAspTyrThrValProSerSerAlaValAspGlyGlyPro 68
     ||||||||         ||||||   ||||||   ::::::||||||
 104 CAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAAGGCGGAGGA 153

  69 AlaAlaGluLeu......AspCysValIleValGlyAlaGlyIleSerGl 83
     ::::::   :::      ||||||||||||||||||:::|||||||||||
 154 GGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGGTATTAGTGG 203

  83 yLeuCysIleAlaGlnValMetSerAlaAsnTyr............ProA 96
     ||||||||||||||||::::::::::::::::::            ||||
 204 TCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATGCTGCTCCGA 253

  96 snLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThr 112
     |||||:::||||||||||||:::||||||:::||||||||||||   |||
 254 ATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAACATTATCACT 303

 113 ValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnPr 129
        |||::::::|||:::||||||||||||||||||||||||||||||||
 304 CGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAGTTTTCAACC 353

 129 oSerAspProMetLeuThrMetAlaValAspCysGlyLeuLysAspAspL 146
     ||||||||||||||||||||||:::||||||   ||||||||||||||||
 354 GTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGAAGGATGATT 403

 146 euValLeuGlyAspProAsnAlaProArgPheValLeuTrpLysGlyLys 162
     |||||||||||||||||:::|||||||||||||||||||||:::||||||
 404 TGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGGAATGGGAAA 453

 163 LeuArgProValProSerLysLeuThrAspLeuProPhePheAspLeuMe 179
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 454 TTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTTTGATTTGAT 503

 179 tSerIleProGlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgP 196
     |||||||   ||||||:::|||||||||||||||   :::|||:::||||
 504 GAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTGGCATTCGAC 553

 196 roSerProProGlyHisGluGluSerValGluGlnPheValArgArgAsn 212
     ||||||||||||||:::|||||||||||||||:::|||||||||||||||
 554 CGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTACGGCGTAAC 603

 213 LeuGlyGlyGluValPheGluArgLeuIleGluProPheCysSerGlyVa 229
     ||||||   |||||||||||||||||||||||||||||||||||||||||
 604 CTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTGTTCAGGTGT 653

 229 lTyrValGlyAspProSerLysLeuSerMetLysAlaAlaPheGlyLysV 246
     ||||:::|||||||||||||||||||||||||||||||||||||||||||
 654 TTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGTTTGGGAAGG 703

 246 alTrpLysLeuGluGluThrGlyGlySerIleIleGlyGlyThrPheLys 262
     ||||||||||||||::::::||||||||||||||||||||||||||||||
 704 TTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGTACTTTTAAG 753

 263 AlaIleLysGluArgSerSerThrProLysAlaProArgAspProArgLe 279
     ||||||:::||||||:::::::::|||||||||   ||||||||||||||
 754 GCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGACCCGCGCCT 803

 279 uProLysProLysGlyGlnThrValGlySerPheArgLysGlyLeuArgM 296
     ||||||||||:::|||||||||||||||||||||||||||||||||||||
 804 GCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGGGACTTCGAA 853

 296 etLeuProAspAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSer 312
     ||||||||:::|||||||||||||||||||||||||||:::|||||||||
 854 TGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTTAAGTTGTCT 903

 313 TrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTyrHisLeuTh 329
     ||||||||||||:::|||||||||   ||||||:::|||||||||:::|||||
 904 TGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATACAACTTAAC 953

 329 rTyrGluThrProGluGlyValValSerLeuGlnSerArgSerIleValM 346
     |||||||||||||:::|||:::||||||:::||||||:::|||:::||||
 954 ATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAAGTGTTGTAA 1003

 346 etThrValProSerTyrValAlaSerAsnIleLeuArgProLeuSerVal 362
     ||||||||||||||:::|||||||||::::::|||||||||||||||
1004 TGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCTCTTTCTGAA 1053

 363 AlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProValGlyAlaVa 379
     :::||||||:::|||||||||::::::|||||||||||||||:::|||||
1054 TCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGTTGCAGCAGT 1103

 379 lThrIleSerTyrProGlnGluAlaIleArgAspGluArgLeuValAspG 396
     |:::||||||||||||:::||||||||||||   |||   |||:::||||
1104 ATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTTTGATAGATG 1153

 396 lyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGlu 412
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1154 GTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAAGGAGTTGAA 1203

 413 ThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnArgAlaProLy 429
     ||||||||||||||||||||||||||||||||||||||||||||||||
1204 ACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCGCGCACCGCC 1253

 429 sGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluI 446
      ||||||:::||||||||||||||||||||||||:::   |||      |
1254 CGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAAACACCGGAA 1303

 446 leLeuSerLysThrGluSerGlnLeuValGluValValAspArgAspLeu 462
     |||||||||||:::|||::::::|||||||||:::|||||||||||||||
1304 TTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGACAGAGATTTG 1353

 463 ArgLysMetLeuIleLysProLysAlaGlnAspProLeuValValGlyVa 479
     |||||||||||||||||||||::::::   |||||||||   :::|||||
1354 AGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAAATTAGGAGT 1403

 479 lArgValTrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAspT 496
     |||||||||||||||||||||||||||||||||||||||||||:::|||
1404 TAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTCACTTTGATA 1453

 496 hrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeu 512
       |||:::|||||||||::::::::::::::::::|||   |||||||||
1454 TCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTACGAAGGGCTA 1503

 513 PheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyArgCysValGl 529
     |||||||||||||||||||||:::||||||||||||||||||||||||||
1504 TTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCGGTGTGTAGA 1553

 529 uGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuSerArgTyrA 546
     |||||||||||||:::|||   ||||||::::::|||:::||||||||||
1554 AGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGTCACGGTACG 1603

 546 laTyrLys 548
     ||||||||
1604 CTTACAAG 1611

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76571

seq_documentation_block:
ID   AAF76571 standard; cDNA; 1719 BP.
XX
AC   AAF76571;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Arabidopsis protoporphyrinogen oxidase coding sequence SEQ ID NO: 1.
XX
KW   Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
```

```
KW   soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200112825-A1.
XX
PD   22-FEB-2001.
XX
PF   30-JUN-2000; 2000WO-EP06127.
XX
PR   13-AUG-1999;   99US-0373691.
XX
PA   (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX
PI   Johnson MA,  Volrath SL,  Heifetz PB,  Law MD;
XX
DR   WPI; 2001-234914/24.
DR   P-PSDB; AAB72901.
XX
PT   Plant DNA molecules encoding herbicide-tolerant forms of
PT   protoporphyrinogen oxidase which are useful for rationally designing
PT   new inhibitory herbicides and for producing herbicide-tolerant
PT   transgenic plants and seeds -
XX
PS   Claim 2; Page 133-137; 228pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
CC   sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
CC   oxidase (protox) enzyme. Examples of these mutants are shown in
CC   AAB72920-AAB72926. They are useful as they enable the production of
CC   herbicide-tolerant plants and seeds. The present sequence is a protox
CC   coding sequence.
XX
SQ   Sequence 1719 BP; 460 A; 351 C; 432 G; 476 T; 0 other;


alignment_scores:
           Quality: 2068.00            Length:    539
             Ratio:  4.220               Gaps:      4
 Percent Similarity:  90.909   Percent Identity: 74.954

alignment_block:
 US-09-508-418-2 x AAF76571   ..

 Align seg 1/1  to: AAF76571  from: 1  to: 1719

   19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
      |||      |||::::::|||:::   |||:::         :::   ||
   28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

   35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
      |:::||||||   |||   ::::::   |||   :::         |||
   78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

   49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
        |||||||||||||||         ||||||   ||||||   ::::::
  125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

   66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
      ||||||   ::::::   :::      ||||||||||||||||||:::||
  175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

   80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
      |||||||||||||||||||||||::::::::::::::::::
  225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274

   95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
           ||||||||:::||||||||||||:::||||||:::||||||||||
  275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324

  110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
      |||   |||   |||::::::|||:::|||||||||||||||||||||||

  325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

  126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
      |||||||||||||||||||||||||||||||:::||||||   |||||||
  375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

  143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
      ||||||||||||||||||||||||||:::|||||||||||||||||||||
  425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

  160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
      :::|||||||||||||||||||||||||||||||||||||||||||||||
  475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524

  176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
      ||||||||||||||||   ||||||:::|||||||||||||||   :::|
  525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

  193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
      ||:::||||||||||||||||||:::|||||||||||||||:::||||||
  575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

  210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
      |||||||||||||||   ||||||||||||||||||||||||||||||||
  625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674

  226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
      |||||||||||||:::||||||||||||||||||||||||||||||||||
  675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

  243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
      |||||||||||||||||||||||::::::|||||||||||||||||||||
  725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

  260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
      |||||||||||||||:::||||||:::::::::|||||||||   |||||
  775 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

  276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
      |||||||||||||||||||:::||||||||||||   |||||||||||||
  825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGTTTCTTTCAGGAAGG 874

  293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
      |||||||||||||||||:::|||||||||||||||||||||||||||:::
  875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

  310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
      |||||||||||||||||||||:::|||||||||   |||:::||||||||
  925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

  326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
      |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
  975 CAACTTAACATATGAGACTCCAGATGGTTAGTTTCCGTGCAGAGCAAAA 1024

  343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
      ||:::||||||||||||||||||:::|||||||||::::::|||||||||
 1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

  360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
      ||||||   :::||||||:::|||||||||::::::||||||||||||||
 1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

  376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
      |:::||||||:::||||||||||||:::||||||||||||   |||   |
 1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

  393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
      ||:::|||||||||||||||||||||||||||||||||||||||||||||
 1175 TGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAA 1224

  410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274
```

```
     426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
         |||||||   ||||||:::||||||||||||||||||||||||:::   |
    1275 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

     443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
         ||      ||||||||||||:::|||::::::|||||||||:::||||||
    1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

     460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
         ||||||||||||||||||||||||||||||::::::   |||||||||
    1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

     476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
          :::||||||||||||||||||||||||||||||||||||||||||||||
    1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

     493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
         ||:::|||   |||:::|||||||||::::::::::::::::::|||
    1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

     510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
         ||||||||||||||||||||||||||||||:::|||||||||||||||||
    1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

     526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
         ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
    1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

     543 erArgTyrAlaTyrLys 548
         |||||||||||||||||
    1625 CACGGTACGCTTACAAG 1641

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ45271

seq_documentation_block:
ID   AAZ45271 standard; DNA; 1632 BP.
XX
AC   AAZ45271;
XX
DT   27-MAR-2000  (first entry)
XX
DE   DNA encoding the chloroplast-type soybean protoporphyrin IX oxidase.
XX
KW   Weed control; weed activity; plant; protoporphyrin IX oxidase; PPO;
KW   herbicide; chlormethoxynil; acifluorfen; protoporphyrin X; soybean; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1632
FT                   /*tag=  a
FT                   /product= "protoporphyrin IX oxidase"
XX
PN   EP953646-A2.
XX
PD   03-NOV-1999.
XX
PF   30-APR-1999;   99EP-0108463.
XX
PR   30-APR-1998;   98JP-0120553.
PR   02-OCT-1998;   98JP-0281127.
PR   20-NOV-1998;   98JP-0330981.
PR   02-MAR-1999;   99JP-0054730.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI   Nakajima H,  Nagasawa A;
XX
DR   WPI; 2000-088762/08.
DR   P-PSDB; AAY54100.
XX
PT   Method of making plants resistant to weed control compounds by
PT   introducing gene encoding protein into plant cell -
XX
PS   Example 3; Page 60-64; 119pp; English.
XX
CC   The present sequence encodes the soybean protoporphyrin IX oxidase
CC   (PPO) gene. The PPO gene and protien are used in the course of the
CC   invention. The specification describes a method of making plants
CC   resistant to weed control compounds. The method comprises introducing
CC   a gene encoding a protein into a plant cell, where the protein has a
CC   specific affinity for a substance which is concerned with the weed
CC   activity of a weed control compound, e.g. protoporphyrin IX; has no
CC   ability to modify this protein; and is virtually free from framework
CC   regions in an immunoglobulin; and expressing the gene. The gene of the
CC   invention is used for producing a plant resistant to specified compounds,
CC   such as protoporphyrin IX oxidase (PPO) inhibitory-type herbicides, e.g.
CC   chlormethoxynil, acifluorfen, etc.. When PPO is inhibited,
CC   protoporphyrin IX (the substrate of PPO) accumulates in the plant cells,
CC   and is metabolised to form protoporphyrin X, followed by formation of
CC   active oxygen in the presence of protoporphyrin X and light, which
CC   damages cell functions.
XX
SQ   Sequence 1632 BP; 404 A; 387 C; 409 G; 432 T; 0 other;


alignment_scores:
            Quality: 2062.50              Length:    550
              Ratio:  4.218                 Gaps:      7
 Percent Similarity:  88.909   Percent Identity: 73.273

alignment_block:
 US-09-508-418-2 x AAZ45271    ..

 Align seg 1/1  to: AAZ45271  from: 1  to: 1632

       3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
         ::::::|||:::         ||||||:::|||   |||   |||   ||
      48 TCTTCGCCCCTCCCTCCATTCCCCAACCTCTTTCTTCACCTCTCCCACTC 97

      19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerI 36
         |:::                              ||||||
      98 GAAA..............................ATTCCC.........T 108

      36 le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
            |||:::   |||   :::                     ||||||||
     109 CGCTCTCGCCCTAACCCTATT..................CTACGCTGCTC 140

      52 rValAlaLysAspTyrThrVal...ProSerSerAlaValAspGlyGlyP 68
         |:::|||::::::   |||:::   |||   ::::::   |||::::::|
     141 CATTGCGGAGGAATCCACCGCGTCTCCGCCCAAAACCAGAGACTCCGCCC 190

      68 roAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
         ||         :::|||||||||:::||||||:::|||:::|||||||||
     191 CC.........GTGGACTGCGTCGTCGTCGGCGGAGGCGTCAGCGGCCTC 231

      85 CysIleAlaGlnValMetSerAlaAsnTyr......ProAsnLeuMetVa 99
         ||||||||||||::::::::::::::::::         |||::::::||
     232 TGCATCGCCCAGGCCCTCGCCACCAAACACGCCAATGCCAACGTCGTCGT 281

      99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgA 116
         |||||||||||||||||||:::||||||||||||||||||:::|||||||
     282 CACGGAGGCCCGAGACCGCGTCGGCGGCAACATCACCACGATGGAGAGGG 331

     116 spGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     332 ACGGATACCTCTGGGAAGAAGGCCCCAACAGCTTCCAGCCTTCTGATCCA 381

     133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGl 149
         ||||||||||||:::||||||   ||||||||||||:::|||||||||||
     382 ATGCTCACCATGGTGGTGGACAGTGGTTTAAAGGATGAGCTTGTTTTGGG 431

     149 yAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
         |||||||:::|||||||||||||||||||||:::   |||||||||||||
     432 GGATCCTGATGCACCTCGGTTTGTGTTGTGGAACAGGAAGTTGAGGCCGG 481
```

```
 166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
     |||||:::|||||||||||||||||||||||||||||||||||||||
 482 TGCCCGGGAAGCTGACTGATTTGCCTTTCTTTGACTTGATGAGCATTGGT 531

 183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
     ||||||:::|||||||||||||||   :::|||:::||||||   |||||
 532 GGCAAAATCAGGGCTGGCTTTGGTGCGCTTGGAATTCGGCCTCCTCCTCC 581

 199 oGlyHisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyG 216
     ||||||||||||||||||||||:::|||||||||||||||||||||   |
 582 AGGTCATGAGGAATCGGTTGAAGAGTTTGTTCGTCGGAACCTTGGTGATG 631

 216 luValPheGluArgLeuIleGluProPheCysSerGlyValTyrValGly 232
     ||||||||||||||||||||||||||||||||||||||||||||:::|||
 632 AGGTTTTTGAACGGTTGATAGAGCCTTTTTGTTCAGGGGTCTATGCAGGC 681

 233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLe 249
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 682 GATCCTTCAAAATTAAGTATGAAAGCAGCATTCGGGAAAGTTTGGAAGCT 731

 249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
     ||||::::::||||||||||||||||||||||||||||||||||||:::|
 732 GGAAAAAAATGGTGGTAGCATTATTGGTGGAACTTTCAAAGCAATACAAG 781

 266 luArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
     |||||:::::::::   |||   |||||||||||||||||||||||||||
 782 AGAGAAATGGAGCTTCAAAACCACCTCGAGATCCGCGTCTGCCAAAACCA 831

 283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
     ||||||||||||||||||||||||||||||||||||   |||||||||||
 832 AAAGGTCAGACTGTTGGATCTTTCCGGAAGGGACTTACCATGTTGCCTGA 881

 299 pAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
     ||||||||||||||||||||||:::|||:::|||||||||||||||||||
 882 TGCAATTTCTGCCAGACTAGGCAACAAAGTAAAGTTATCTTGGAAGCTTT 931

 316 erSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThr 332
     ||||||||:::|||   ::::::|||   |||   |||||||||||||||
 932 CAAGTATTAGTAAACTGGATAGTGGAGAGTACAGTTTGACATATGAAACA 981

 333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetThrValPr 349
     ||||||||||||||||||||||||   :::::::::|||:::|||:::||
 982 CCAGAAGGAGTGGTTTCTTTGCAGTGCAAAACTGTTGTCCTGACCATTCC 1031

 349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaA 366
     ||||||||||||||||::::::|||||||||||||||:::||||||||||
1032 TTCCTATGTTGCTAGTACATTGCTGCGTCCTCTGTCTGCTGCTGCTGCAG 1081

 366 spAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSer 382
     |||||||||||:::||||||||||||||||||:::||||||:::||||||
1082 ATGCACTTTCAAAGTTTTATTACCCTCCAGTTGCTGCAGTTTCCATATCC 1131

 383 TyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLy 399
     ||||||:::||||||||||||:::|||   |||:::||||||||||||||
1132 TATCCAAAAGAAGCTATTAGATCAGAATGCTTGATAGATGGTGAGTTGAA 1181

 399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyT 416
     |||||||||||||||||||||||||:::||||||||||||||||||||||
1182 GGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAACATTAGGAA 1231

 416 hrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
     ||||||||||||||||||||||||||||||||||||||   |||||||||
1232 CTATATACAGCTCATCACTATTCCCCAACCGAGCACCACCTGGAAGGGTT 1281

 433 LeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLy 449
     |||||||||||||||||||||||||||   |||      |||||||||||
1282 CTACTCTTGAATTACATTGGAGGAGCAACTAATACTGGAATTTTATCGAA 1331

 449 sThrGluSerGlnLeuValGluValValAspArgAspLeuArgLysMetL 466
     ||||:::|||:::|||||||||:::|||||||||||||||||||||:::|
1332 GACGGACAGTGAACTTGTGGAAACAGTTGATCGAGATTTGAGGAAAATCC 1381
```

```
 466 euIleLysProLysAlaGlnAspProLeuValValGlyValArgValTrp 482
     |||||:::|||:::||||||||||||:::|||||||||||||||:::|||
1382 TTATAAACCCAAATGCCCAGGATCCATTTGTAGTGGGGTGAGACTGTGG 1431

 483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
     |||||||||||||||||||||||||||||||||||||||   |||:::::
1432 CCTCAAGCTATTCCACAGTTCTTAGTTGGCCATCTTGATCTTCTAGATGT 1481

 499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
     :|||||||||:::::::::::::::|||:::|||||||||||||||||||
1482 TGCTAAAGCTTCTATCAGAAATACTGGGTTTGAAGGGCTCTTCCTTGGGG 1531

 516 lyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1532 GTAATTATGTGTCTGGTGTTGCCTTGGGACGATGCGTTGAGGGAGCCTAT 1581

 533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
     |||||||||:::||||||:::   ||||||::::::   :::||||||
1582 GAGGTAGCAGCTGAAGTAAACGATTTTCTCACAAATAGAGTGTACAAA 1629

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAV04309

seq_documentation_block:
ID   AAV04309 standard; cDNA; 1847 BP.
XX
AC   AAV04309;
XX
DT   20-APR-1998  (first entry)
XX
DE   Soybean protox-1 cDNA.
XX
KW   Protoporphyrinogen oxidase-1; protox-1; promoter; soybean;
KW   herbicide resistance; breeding programme; probe; gene isolation;
KW   genomic mapping; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             55..1686
FT                   /*tag= a
FT                   /product= protox-1
XX
PN   WO9732028-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US03343.
XX
PR   21-JUN-1996;   96US-0020003.
PR   28-FEB-1996;   96US-0012705.
PR   28-FEB-1996;   96US-0013612.
XX
PA   (NOVS ) NOVARTIS AG.
XX
PI   Johnson MA,  Volrath SL,  Ward ER;
XX
DR   WPI; 1997-489209/45.
DR   P-PSDB; AAW41608.
XX
PT   DNA containing a plant proto-porphyrinogen oxidase gene promoter -
PT   optionally linked to a heterologous gene, especially to express
PT   herbicide-resistant enzymes, and plants containing such constructs
XX
PS   Disclosure; Pages 68-71; 114pp; English.
XX
CC   The present sequence encodes soybean protoporphyrinogen oxidase-1
CC   (protox-1).
CC   The protox-1 promoter can be used to express herbicide resistant
CC   enzymes, specifically protox, i.e. a plant tissue, plant or progeny
CC   containing a chimeric gene of the promoter and a heterologous
CC   coding sequence. The plant can also be used in breeding programmes.
CC   Also hybridising fragments of the protox coding sequence can be
CC   used as probes, e.g. to isolate related genes or for genomic
```

```
CC   mapping.
XX
SQ   Sequence 1847 BP; 473 A; 416 C; 453 G; 505 T; 0 other;


alignment_scores:
           Quality: 2062.50            Length:     550
             Ratio:  4.218               Gaps:       7
 Percent Similarity:  88.909   Percent Identity: 73.273

alignment_block:
 US-09-508-418-2 x AAV04309   ..

 Align seg 1/1  to: AAV04309  from: 1  to: 1847

       3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
         ::::::|||:::         ||||||:::|||   |||   |||   ||
     102 TCTTCGCCCCTCCCTCCATTCCCCAACCTCTTTCTTCACCTCTCCCACTC 151

      19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerI 36
         |:::                              ||||||
     152 GAAA..............................ATTCCC.........T 162

      36 le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
            |||:::   |||   :::                     ||||||||
     163 CGCTCTCGCCCTAACCCTATT..................CTACGCTGCTC 194

      52 rValAlaLysAspTyrThrVal...ProSerSerAlaValAspGlyGlyP 68
         |:::|||::::::   |||:::   |||   ::::::   |||::::::|
     195 CATTGCGGAGGAATCCACCGCGTCTCCGCCCAAAACCAGAGACTCCGCCC 244

      68 roAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
         ||         :::|||||||||:::||||||:::|||:::|||||||||
     245 CC.........GTGGACTGCGTCGTCGTCGGCGGAGGCGTCAGCGGCCTC 285

      85 CysIleAlaGlnValMetSerAlaAsnTyr......ProAsnLeuMetVa 99
         ||||||||||||::::::::::::::::::         |||::::::||
     286 TGCATCGCCCAGGCCCTCGCCACCAAACACGCCAATGCCAACGTCGTCGT 335

      99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgA 116
         |||||||||||||||||||:::||||||||||||||||||:::|||||||
     336 CACGGAGGCCCGAGACCGCGTCGGCGGCAACATCACCACGATGGAGAGGG 385

     116 spGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     386 ACGGATACCTCTGGGAAGAAGGCCCCAACAGCTTCCAGCCTTCTGATCCA 435

     133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGl 149
         ||||||||||||:::||||||   ||||||||||||:::|||||||||||
     436 ATGCTCACCATGGTGGTGGACAGTGGTTTAAAGGATGAGCTTGTTTTGGG 485

     149 yAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
         |||||||:::|||||||||||||||||||||:::   |||||||||||||
     486 GGATCCTGATGCACCTCGGTTTGTGTTGTGGAACAGGAAGTTGAGGCCGG 535

     166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
         |||||:::|||||||||||||||||||||||||||||||||||||||
     536 TGCCCGGGAAGCTGACTGATTTGCCTTTCTTTGACTTGATGAGCATTGGT 585

     183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
         ||||||:::|||||||||||||||   :::|||:::||||||   |||||
     586 GGCAAAATCAGGGCTGGCTTTGGTGCGCTTGGAATTCGGCCTCCTCCTCC 635

     199 oGlyHisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyG 216
         ||||||||||||||||||||||:::|||||||||||||||||||||   |
     636 AGGTCATGAGGAATCGGTTGAAGAGTTTGTTCGTCGGAACCTTGGTGATG 685

     216 luValPheGluArgLeuIleGluProPheCysSerGlyValTyrValGly 232
         ||||||||||||||||||||||||||||||||||||||||||||:::|||
     686 AGGTTTTTGAACGGTTGATAGAGCCTTTTTGTTCAGGGGTCTATGCAGGC 735

     233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLe 249
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     736 GATCCTTCAAAATTAAGTATGAAAGCAGCATTCGGGAAAGTTTGGAAGCT 785

     249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
         ||||::::::||||||||||||||||||||||||||||||||||||:::|
     786 GGAAAAAAATGGTGGTAGCATTATTGGTGGAACTTTCAAAGCAATACAAG 835

     266 luArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
         |||||:::::::::   |||   |||||||||||||||||||||||||||
     836 AGAGAAATGGAGCTTCAAAACCACCTCGAGATCCGCGTCTGCCAAAACCA 885

     283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
         ||||||||||||||||||||||||||||||||||||   |||||||||||
     886 AAAGGTCAGACTGTTGGATCTTTCCGGAAGGGACTTACCATGTTGCCTGA 935

     299 pAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
         ||||||||||||||||||||||:::|||:::|||||||||||||||||||
     936 TGCAATTTCTGCCAGACTAGGCAACAAAGTAAAGTTATCTTGGAAGCTTT 985

     316 erSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThr 332
         ||||||||:::|||   ::::::|||   |||   |||||||||||||||
     986 CAAGTATTAGTAAACTGGATAGTGGAGAGTACAGTTTGACATATGAAACA 1035

     333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetThrValPr 349
         ||||||||||||||||||||||||   :::::::::|||:::|||:::||
    1036 CCAGAAGGAGTGGTTTCTTTGCAGTGCAAAACTGTTGTCCTGACCATTCC 1085

     349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaA 366
         ||||||||||||||||::::::|||||||||||||||:::||||||||||
    1086 TTCCTATGTTGCTAGTACATTGCTGCGTCCTCTGTCTGCTGCTGCTGCAG 1135

     366 spAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSer 382
         |||||||||||:::||||||||||||||||||:::||||||:::||||||
    1136 ATGCACTTTCAAAGTTTATTACCCTCCAGTTGCTGCAGTTTCCATATCC 1185

     383 TyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLy 399
         ||||||:::||||||||||||:::|||   |||:::||||||||||||||
    1186 TATCCAAAAGAAGCTATTAGATCAGAATGCTTGATAGATGGTGAGTTGAA 1235

     399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyT 416
         |||||||||||||||||||||||||:::||||||||||||||||||||||
    1236 GGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAACATTAGGAA 1285

     416 hrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
         ||||||||||||||||||||||||||||||||||||||   |||||||||
    1286 CTATATACAGCTCATCACTATTCCCCAACCGAGCACCACCTGGAAGGGTT 1335

     433 LeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLy 449
         |||||||||||||||||||||||||||   |||      |||||||||||
    1336 CTACTCTTGAATTACATTGGAGGAGCAACTAATACTGGAATTTTATCGAA 1385

     449 sThrGluSerGlnLeuValGluValValAspArgAspLeuArgLysMetL 466
         ||||:::|||:::|||||||||:::|||||||||||||||||||||:::|
    1386 GACGGACAGTGAACTTGTGGAAACAGTTGATCGAGATTTGAGGAAAATCC 1435

     466 euIleLysProLysAlaGlnAspProLeuValValGlyValArgValTrp 482
         |||||:::|||:::||||||||||||:::|||||||||||||||:::|||
    1436 TTATAAACCCAAATGCCCAGGATCCATTTGTAGTGGGGGTGAGACTGTGG 1485

     483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
         |||||||||||||||||||||||||||||||||||||||   |||:::::
    1486 CCTCAAGCTATTCCACAGTTCTTAGTTGGCCATCTTGATCTTCTAGATGT 1535

     499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
         :|||||||||:::::::::::::::|||:::|||||||||||||||||||
    1536 TGCTAAAGCTTCTATCAGAAATACTGGGTTTGAAGGGCTCTTCCTTGGGG 1585

     516 lyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1586 GTAATTATGTGTCTGGTGTTGCCTTGGGACGATGCGTTGAGGGAGCCTAT 1635

     533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
         |||||||||:::||||||:::   ||||||::::::   :::||||||
    1636 GAGGTAGCAGCTGAAGTAAACGATTTCTCACAAATAGAGTGTACAAA 1683
```

```
seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT86122

seq_documentation_block:
ID   AAT86122 standard; cDNA; 1847 BP.
XX
AC   AAT86122;
XX
DT   01-MAR-1998  (first entry)
XX
DE   Soybean protoporphyrinogen oxidase (protox-1) cDNA clone pWDC-12.
XX
KW   Protox-1; protoporphyrinogen oxidase; inhibitor; soybean;
KW   herbicide tolerance; herbicide resistance; transgenic plant; ss.
XX
OS   Glycine max var. Williams 82.
XX
FH   Key             Location/Qualifiers
FT   CDS             55..1686
FT                   /*tag=  a
XX
PN   WO9732011-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US03313.
XX
PR   21-JUN-1996;   96US-0020003.
PR   28-FEB-1996;   96US-0012705.
PR   28-FEB-1996;   96US-0013612.
XX
PA   (NOVS ) NOVARTIS AG.
XX
PI   Heifetz PB,  Johnson MA,  Potter SL,  Volrath SL,  Ward ER;
XX
DR   WPI; 1997-448683/41.
DR   P-PSDB; AAW25739.
XX
PT   New DNA encoding plant protoporphyrinogen oxidase enzyme  - and
PT   herbicide resistant mutants, useful to prepare plants resistant to
PT   herbicide which therefore kills undesired vegetation only
XX
PS   Claim 6; Page 127-130; 196pp; English.
XX
CC   This cDNA clone codes for soybean protoporphyrinogen oxidase
CC   (Protox-1, see AAW25739).  It was isolated from an epicotyl cDNA
CC   library using Arabidopsis Protox-1 cDNA (see AAT86129) as probe.
CC   Soybean Protox-1 in pBluescript SK vector has been deposited as
CC   pWDC-12 (NRRL AAB21516).  The isolated cDNA can be mutated so that
CC   it encodes a modified protox enzyme that is resistant to protox
CC   inhibitors and hence is herbicide tolerant.  Plants, especially
CC   crop plants, may be engineered for resistance to protox inhibitors
CC   via mutation of the native protox gene to a resistant form, or they
CC   may be transformed with a gene encoding an inhibitor-resistant form
CC   of a plant protox enzyme, including claimed forms from wheat,
CC   soybean, cotton, sugarbeet, oilseed rape, rice and sorghum (see
CC   AAW25738-45).  Application of herbicide therefore kills undesired
CC   vegetation only.  Plant protox nucleic acids can also be used as
CC   probes and PCR primers, as selectable markers in plant cell
CC   transformation methods, and for recombinant production of protox
CC   enzymes in host cells.
XX
SQ   Sequence 1847 BP; 473 A; 416 C; 453 G; 505 T; 0 other;


alignment_scores:
           Quality: 2062.50            Length:    550
             Ratio: 4.218                Gaps:      7
 Percent Similarity:  88.909   Percent Identity: 73.273

alignment_block:
 US-09-508-418-2 x AAT86122   ..
```

```
Align seg 1/1  to: AAT86122  from: 1  to: 1847

    3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
      ::::::|||:::         ||||||:::|||   |||   |||   ||
  102 TCTTCGCCCCTCCCTCCATTCCCCAACCTCTTTCTTCACCTCTCCCACTC 151

   19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerI 36
      |:::                              ||||||
  152 GAAA..............................ATTCCC.........T 162

   36 le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
         |||:::   |||   :::                     ||||||||
  163 CGCTCTCGCCCTAACCCTATT..................CTACGCTGCTC 194

   52 rValAlaLysAspTyrThrVal...ProSerSerAlaValAspGlyGlyP 68
      |:::|||::::::   |||:::   |||   ::::::   |||::::::|
  195 CATTGCGGAGGAATCCACCGCGTCTCCGCCCAAAACCAGAGACTCCGCCC 244

   68 roAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
      ||         :::|||||||||:::||||||:::|||:::|||||||||
  245 CC.........GTGGACTGCGTCGTCGTCGGCGGAGGCGTCAGCGGCCTC 285

   85 CysIleAlaGlnValMetSerAlaAsnTyr......ProAsnLeuMetVa 99
      ||||||||||||::::::::::::::::::         |||::::::||
  286 TGCATCGCCCAGGCCCTCGCCACCAAACACGCCAATGCCAACGTCGTCGT 335

   99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgA 116
      |||||||||||||||||||:::||||||||||||||||||:::|||||||
  336 CACGGAGGCCCGAGACCGCGTCGGCGGCAACATCACCACGATGGAGAGGG 385

  116 spGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  386 ACGGATACCTCTGGGAAGAAGGCCCCAACAGCTTCCAGCCTTCTGATCCA 435

  133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGl 149
      |||||||||||||:::|||||   ||||||||||||:::|||||||||||
  436 ATGCTCACCATGGTGGTGGACAGTGGTTTAAAGGATGAGCTTGTTTTGGG 485

  149 yAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
      |||||||:::|||||||||||||||||||||:::   |||||||||||||
  486 GGATCCTGATGCACCTCGGTTTGTGTTGTGGAACAGGAAGTTGAGGCCGG 535

  166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
      |||||:::|||||||||||||||||||||||||||||||||||||||
  536 TGCCCGGGAAGCTGACTGATTTGCCTTTCTTTGACTTGATGAGCATTGGT 585

  183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
      ||||||:::|||||||||||||||   :::|||:::||||||   |||||
  586 GGCAAAATCAGGGCTGGCTTTGGTGCGCTTGGAATTCGGCCTCCTCCTCC 635

  199 oGlyHisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyG 216
      ||||||||||||||||||||||:::|||||||||||||||||||||   |
  636 AGGTCATGAGGAATCGGTTGAAGAGTTTGTTCGTCGGAACCTTGGTGATG 685

  216 luValPheGluArgLeuIleGluProPheCysSerGlyValTyrValGly 232
      ||||||||||||||||||||||||||||||||||||||||||||:::|||
  686 AGGTTTTTGAACGGTTGATAGAGCCTTTTTGTTCAGGGGTCTATGCAGGC 735

  233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLe 249
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  736 GATCCTTCAAAATTAAGTATGAAAGCAGCATTCGGGAAAGTTTGGAAGCT 785

  249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
      ||||::::::||||||||||||||||||||||||||||||||||||:::|
  786 GGAAAAAAATGGTGGTAGCATTATTGGTGGAACTTTCAAAGCAATACAAG 835

  266 luArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
      |||||:::::::::   |||   |||||||||||||||||||||||||||
  836 AGAGAAATGGAGCTTCAAAACCACCTCGAGATCCGCGTCTGCCAAAACCA 885

  283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
      ||||||||||||||||||||||||||||||||||||   |||||||||||
  886 AAAGGTCAGACTGTTGGATCTTTCCGGAAGGGACTTACCATGTTGCCTGA 935
```

```
  299 pAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
      ||||||||||||||||||||||:::|||:::|||||||||||||||||||
  936 TGCAATTTCTGCCAGACTAGGCAACAAAGTAAAGTTATCTTGGAAGCTTT 985

  316 erSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThr 332
      ||||||||:::|||   ::::::|||   |||   |||||||||||||||
  986 CAAGTATTAGTAAACTGGATAGTGGAGAGTACAGTTTGACATATGAAACA 1035

  333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetThrValPr 349
      ||||||||||||||||||||||||   :::::::::|||:::|||:::||
 1036 CCAGAAGGAGTGGTTTCTTTGCAGTGCAAAACTGTTGTCCTGACCATTCC 1085

  349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaA 366
      ||||||||||||||||::::::|||||||||||||||:::||||||||||
 1086 TTCCTATGTTGCTAGTACATTGCTGCGTCCTCTGTCTGCTGCTGCTGCAG 1135

  366 spAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSer 382
      |||||||||||:::||||||||||||||||||:::||||||:::||||||
 1136 ATGCACTTTCAAAGTTTTATTACCCTCCAGTTGCTGCAGTTTCCATATCC 1185

  383 TyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLy 399
      ||||||:::||||||||||||:::|||   |||:::||||||||||||||
 1186 TATCCAAAAGAAGCTATTAGATCAGAATGCTTGATAGATGGTGAGTTGAA 1235

  399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyT 416
      |||||||||||||||||||||||||:::||||||||||||||||||||||
 1236 GGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAACATTAGGAA 1285

  416 hrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
      ||||||||||||||||||||||||||||||||||||||   |||||||||
 1286 CTATATACAGCTCATCACTATTCCCCAACCGAGCACCACCTGGAAGGGTT 1335

  433 LeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLy 449
      |||||||||||||||||||||||||||   |||      |||||||||||
 1336 CTACTCTTGAATTACATTGGAGGAGCAACTAATACTGGAATTTTATCGAA 1385

  449 sThrGluSerGlnLeuValGluValValAspArgAspLeuArgLysMetL 466
      ||||:::|||:::|||||||||:::|||||||||||||||||||||:::|
 1386 GACGGACAGTGAACTTGTGGAAACAGTTGATCGAGATTTGAGGAAAATCC 1435

  466 euIleLysProLysAlaGlnAspProLeuValValGlyValArgValTrp 482
      |||||:::|||:::||||||||||||:::|||||||||||||||:::|||
 1436 TTATAAACCCAAATGCCCAGGATCCATTTGTAGTGGGGGTGAGACTGTGG 1485

  483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
      |||||||||||||||||||||||||||||||||||||||   |||:::::
 1486 CCTCAAGCTATTCCACAGTTCTTAGTTGGCCATCTTGATCTTCTAGATGT 1535

  499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
      :|||||||||:::::::::::::::|||:::|||||||||||||||||||
 1536 TGCTAAAGCTTCTATCAGAAATACTGGGTTTGAAGGGCTCTTCCTTGGGG 1585

  516 lyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1586 GTAATTATGTGTCTGGTGTTGCCTTGGGACGATGCGTTGAGGGAGCCTAT 1635

  533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
      |||||||||:::||||||:::   ||||||::::::   :::||||||
 1636 GAGGTAGCAGCTGAAGTAAACGATTTTCTCACAAATAGAGTGTACAAA 1683

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76576

seq_documentation_block:
ID   AAF76576 standard; cDNA; 1847 BP.
XX
AC   AAF76576;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Soybean protoporphyrinogen oxidase coding sequence SEQ ID NO: 11.
XX
KW   Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
KW   soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein; ss.
XX
OS   Glycine max.
XX
PN   WO200112825-A1.
XX
PD   22-FEB-2001.
XX
PF   30-JUN-2000; 2000WO-EP06127.
XX
PR   13-AUG-1999;   99US-0373691.
XX
PA   (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX
PI   Johnson MA,  Volrath SL,  Heifetz PB,  Law MD;
XX
DR   WPI; 2001-234914/24.
DR   P-PSDB; AAB72906.
XX
PT   Plant DNA molecules encoding herbicide-tolerant forms of
PT   protoporphyrinogen oxidase which are useful for rationally designing
PT   new inhibitory herbicides and for producing herbicide-tolerant
PT   transgenic plants and seeds -
XX
PS   Claim 2; Page 168-172; 228pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
CC   sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
CC   oxidase (protox) enzyme. Examples of these mutants are shown in
CC   AAB72920-AAB72926. They are useful as they enable the production of
CC   herbicide-tolerant plants and seeds. The present sequence is a protox
CC   coding sequence.
XX
SQ   Sequence 1847 BP; 473 A; 416 C; 453 G; 505 T; 0 other;

alignment_scores:
          Quality: 2062.50            Length:    550
            Ratio:  4.218               Gaps:      7
 Percent Similarity:  88.909   Percent Identity: 73.273

alignment_block:
 US-09-508-418-2 x AAF76576   ..

 Align seg 1/1  to: AAF76576  from: 1  to: 1847

        3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
          ::::::|||:::         ||||||:::|||   |||   |||   ||
      102 TCTTCGCCCCTCCCTCCATTCCCCAACCTCTTTCTTCACCTCTCCCACTC 151

       19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerI 36
          |:::                              ||||||
      152 GAAA..............................ATTCCC.........T 162

       36 le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
             |||:::   |||   :::                     ||||||||
      163 CGCTCTCGCCCTAACCCTATT..................CTACGCTGCTC 194

       52 rValAlaLysAspTyrThrVal...ProSerSerAlaValAspGlyGlyP 68
          |:::||||::::::   |||:::   |||   ::::::   |||::::::|
      195 CATTGCGGAGGAATCCACCGCGTCTCCGCCCAAAACCAGAGACTCCGCCC 244

       68 roAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
          ||         :::|||||||||:::||||||:::|||:::|||||||||
      245 CC.........GTGGACTGCGTCGTCGTCGGCGGAGGCGTCAGCGGCCTC 285

       85 CysIleAlaGlnValMetSerAlaAsnTyr......ProAsnLeuMetVa 99
          ||||||||||||::::::::::::::::::         |||::::::||
      286 TGCATCGCCCAGGCCCTCGCCACCAAACACGCCAATGCCAACGTCGTCGT 335

       99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgA 116
          |||||||||||||||||||:::||||||||||||||||||:::|||||||
```

```
 336 CACGGAGGCCCGAGACCGCGTCGGCGGCAACATCACCACGATGGAGAGGG 385

 116 spGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 386 ACGGATACCTCTGGGAAGAAGGCCCCAACAGCTTCCAGCCTTCTGATCCA 435

 133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGl 149
     ||||||||||||:::||||||   ||||||||||||:::|||||||||||
 436 ATGCTCACCATGGTGGTGGACAGTGGTTTAAAGGATGAGCTTGTTTTGGG 485

 149 yAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
     |||||||:::|||||||||||||||||||||:::   |||||||||||||
 486 GGATCCTGATGCACCTCGGTTTGTGTTGTGGAACAGGAAGTTGAGGCCGG 535

 166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
     |||||:::|||||||||||||||||||||||||||||||||||||||
 536 TGCCCGGGAAGCTGACTGATTTGCCTTTCTTTGACTTGATGAGCATTGGT 585

 183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
     ||||||:::|||||||||||||||   :::|||:::||||||   |||||
 586 GGCAAAATCAGGGCTGGCTTTGGTGCGCTTGGAATTCGGCCTCCTCCTCC 635

 199 oGlyHisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyG 216
     ||||||||||||||||||||||:::|||||||||||||||||||||   |
 636 AGGTCATGAGGAATCGGTTGAAGAGTTTGTTCGTCGGAACCTTGGTGATG 685

 216 luValPheGluArgLeuIleGluProPheCysSerGlyValTyrValGly 232
     ||||||||||||||||||||||||||||||||||||||||||||:::|||
 686 AGGTTTTTGAACGGTTGATAGAGCCTTTTTGTTCAGGGGTCTATGCAGGC 735

 233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLe 249
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 736 GATCCTTCAAAATTAAGTATGAAAGCAGCATTCGGGAAAGTTTGGAAGCT 785

 249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
     ||||::::::||||||||||||||||||||||||||||||||||||:::|
 786 GGAAAAAAATGGTGGTAGCATTATTGGTGGAACTTTCAAAGCAATACAAG 835

 266 luArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
     |||||:::::::::   |||   |||||||||||||||||||||||||||
 836 AGAGAAATGGAGCTTCAAAACCACCTCGAGATCCGCGTCTGCCAAAACCA 885

 283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
     ||||||||||||||||||||||||||||||||||||   |||||||||||
 886 AAAGGTCAGACTGTTGGATCTTTCCGGAAGGGACTTACCATGTTGCCTGA 935

 299 pAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
     ||||||||||||||||||||||:::|||:::|||||||||||||||||||
 936 TGCAATTTCTGCCAGACTAGGCAACAAAGTAAAGTTATCTTGGAAGCTTT 985

 316 erSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThr 332
     ||||||||:::|||   ::::::|||   |||   |||||||||||||||
 986 CAAGTATTAGTAAACTGGATAGTGGAGAGTACAGTTTGACATATGAAACA 1035

 333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetThrValPr 349
     ||||||||||||||||||||||||   :::::::::|||:::|||:::||
1036 CCAGAAGGAGTGGTTTCTTTGCAGTGCAAAACTGTTGTCCTGACCATTCC 1085

 349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaA 366
     ||||||||||||||||::::::||||||||||||||||:::||||||||||
1086 TTCCTATGTTGCTAGTACATTGCTGCGTCCTCTGTCTGCTGCTGCTGCAG 1135

 366 spAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSer 382
     |||||||||||:::||||||||||||||||||:::||||||:::||||||
1136 ATGCACTTTCAAAGTTTTATTACCCTCCAGTTGCTGCAGTTTCCATATCC 1185

 383 TyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLy 399
     ||||||:::||||||||||||:::|||   |||:::||||||||||||||
1186 TATCCAAAAGAAGCTATTAGATCAGAATGCTTGATAGATGGTGAGTTGAA 1235

 399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyT 416
     |||||||||||||||||||||||||:::||||||||||||||||||||||
1236 GGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAACATTAGGAA 1285

 416 hrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
     ||||||||||||||||||||||||||||||||||||||   |||||||||
1286 CTATATACAGCTCATCACTATTCCCCAACCGAGCACCACCTGGAAGGGTT 1335

 433 LeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLy 449
     |||||||||||||||||||||||||||   |||      |||||||||||
1336 CTACTCTTGAATTACATTGGAGGAGCAACTAATACTGGAATTTTATCGAA 1385

 449 sThrGluSerGlnLeuValGluValValAspArgAspLeuArgLysMetL 466
     ||||:::|||:::|||||||||:::|||||||||||||||||||||:::|
1386 GACGGACAGTGAACTTGTGGAAACAGTTGATCGAGATTTGAGGAAAATCC 1435

 466 euIleLysProLysAlaGlnAspProLeuValValGlyValArgValTrp 482
     |||||:::|||:::||||||||||||:::|||||||||||||||:::|||
1436 TTATAAACCCAAATGCCCAGGATCCATTTGTAGTGGGGGTGAGACTGTGG 1485

 483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
     |||||||||||||||||||||||||||||||||||||||   |||:::::
1486 CCTCAAGCTATTCCACAGTTCTTAGTTGGCCATCTTGATCTTCTAGATGT 1535

 499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
     :|||||||||:::::::::::::::|||:::|||||||||||||||||||
1536 TGCTAAAGCTTCTATCAGAAATACTGGGTTTGAAGGGCTCTTCCTTGGGG 1585

 516 lyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1586 GTAATTATGTGTCTGGTGTTGCCTTGGGACGATGCGTTGAGGGAGCCTAT 1635

 533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
     |||||||||:::||||||:::   ||||||::::::   :::||||||
1636 GAGGTAGCAGCTGAAGTAAACGATTTTCTCACAAATAGAGTGTACAAA 1683
```

```
OM of: US-09-508-418-2 to: Issued_Patents_NA:*   out_format : pfs

Date: Jul 3, 2001 11:38 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US09508418/runat_03072001_100017_24452/app_query.fasta_1.613
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US09508418_@CGN1_1_66 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-508-418-2
Query length: 548
Database: Issued_Patents_NA:*
Database sequences: 317530
Database length: 92630169
Search time (sec): 94.000000


score_list:
Sequence                  Strd Orig    ZScore   EScore Len  ! Documentation  ..
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-15 + 2113.00 4053.03  3.8e-218  1826  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-15 + 2113.00 4053.03  3.8e-218  1826  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-15 + 2113.00 4053.03  3.8e-218  1826  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-15 + 2113.00 4053.03  3.8e-218  1826  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-472-028A-1 + 2077.00 3984.07  2.6e-214  1719  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-1 + 2077.00 3984.07  2.6e-214  1719  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-1 + 2077.00 3984.07  2.6e-214  1719  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-1 + 2077.00 3984.07  2.6e-214  1719  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-1 + 2077.00 3984.07  2.6e-214  1719  !
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-071-296-1 + 2077.00 3984.07  2.6e-214  1719  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-11 + 2062.50 3955.03  1.1e-212  1847  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-11 + 2062.50 3955.03  1.1e-212  1847  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-11 + 2062.50 3955.03  1.1e-212  1847  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-11 + 2062.50 3955.03  1.1e-212  1847  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-19 + 2006.50 3846.99  1.1e-206  1784  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-19 + 2006.50 3846.99  1.1e-206  1784  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-19 + 2006.50 3846.99  1.1e-206  1784  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-19 + 2006.50 3846.99  1.1e-206  1784  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-9 + 1987.00 3809.01  1.5e-204  1811  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-9 + 1987.00 3809.01  1.5e-204  1811  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-9 + 1987.00 3809.01  1.5e-204  1811  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-9 + 1987.00 3809.01  1.5e-204  1811  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-472-028A-5 + 1978.00 3792.47  1.2e-203  1691  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-5 + 1978.00 3792.47  1.2e-203  1691  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-5 + 1978.00 3792.47  1.2e-203  1691  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-5 + 1978.00 3792.47  1.2e-203  1691  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-5 + 1978.00 3792.47  1.2e-203  1691  !
/cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-071-296-5 + 1978.00 3792.42  1.2e-203  1698  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-17 + 1933.50 3704.65  9.6e-199  1910  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-17 + 1933.50 3704.65  9.6e-199  1910  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-17 + 1933.50 3704.65  9.6e-199  1910  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-17 + 1933.50 3704.65  9.6e-199  1910  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-23 + 1823.00 3492.97  5.9e-187  1590  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-23 + 1823.00 3492.97  5.9e-187  1590  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-23 + 1823.00 3492.97  5.9e-187  1590  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-23 + 1823.00 3492.97  5.9e-187  1590  !
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-21 + 1261.00 2407.52  1.7e-126  1224  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-21 + 1261.00 2407.52  1.7e-126  1224  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-21 + 1261.00 2407.52  1.7e-126  1224  !
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-21 + 1261.00 2407.52  1.7e-126  1224  !
/cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-472-028A-3 +  490.50  910.02  4.4e-43   1738
/cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-3 +  490.50  910.02  4.4e-43   1738
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-3 +  490.50  910.02  4.4e-43   1738
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-3 +  490.50  910.02  4.4e-43   1738
/cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-3 +  490.50  910.02  4.4e-43   1738

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-15

seq_documentation_block:
; Sequence 15, Application US/08808931
; Patent No. 5939602
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5939602artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,931
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1826 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Gossypium hirsutum (cotton)
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-15 (NRRL B-21594)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  31..1647
;      OTHER INFORMATION:  /product= "Cotton protox-1 coding
;      OTHER INFORMATION:  region"
US-08-808-931-15
```

```
alignment_scores:
           Quality:  2113.00              Length:     506
             Ratio:  4.477                  Gaps:       1
 Percent Similarity:  93.281   Percent Identity:  79.249

alignment_block:
 US-09-508-418-2 x US-08-808-931-15    ..

 Align seg 1/1  to: US-08-808-931-15  from: 1  to: 1826

      47 TrpArgThrArgCysSerValAlaLysAspTyrThrValProSerSerAl 63
         ::::::   |||||||||:::|||:::      |||:::   ||||||
     130 TTCAAGCTCCGATGCTCCCTCGCCGAGGGTCCCACGATTTCCTCATCTAA 179

      63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
          :::|||||||||   ::::::      ||||||||||||||||||:::|
     180 AATCGACGGGGGAGAATCATCCATCGCGGATTGCGTCATCGTTGGAGGTG 229

      80 lyIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr...... 94
         ||||||||||||||||||||||||||::::::::::::::::::
     230 GTATCAGTGGACTTTGCATTGCTCAAGCTCTCGCCACCAAGCACCGTGAC 279

      95 ......ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAs 109
                  |||::::::|||||||||||||||||||||:::||||||||
     280 GTCGCTTCCAATGTGATTGTGACGGAGGCCAGAGACCGTGTTGGTGGCAA 329

     109 nIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnS 126
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     330 CATCACTACCGTTGAGAGAGATGGATATCTGTGGGAAGAAGGCCCCAACA 379

     126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
         ||||||||||||||||||||:::||||||||||||||||||   ||||||
     380 GTTTTCAGCCCTCCGATCCTATTCTAACCATGGCCGTGGATAGTGGATTG 429

     143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     430 AAGGACGATTTGGTTTTAGGTGACCCTAATGCACCGCGATTTGTACTATG 479

     159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
         |:::|||||||||||||||||||||||||||   ||||||||||||||||
     480 GGAGGGAAAACTAAGGCCTGTGCCCTCCAAGCCAACCGACTTGCCGTTTT 529

     176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
         |||||||||||||||||   ||||||||||||||||||||||||   |||
     530 TTGATTTGATGAGCATTGCTGGAAAACTTAGGGCTGGGTTCGGGGCTATT 579

     193 GlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVa 209
         |||:::||||||   |||||||||:::|||||||||||||||:::|||||
     580 GGCATTCGGCCTCCCCCTCCGGGTTATGAAGAATCGGTGGAGGAGTTTGT 629

     209 lArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC 226
         ||||||||||||||||:::|||||||||||||||:::|||||||||||||
     630 GCGCCGTAATCTTGGTGCTGAGGTTTTTGAACGCTTTATTGAACCATTTT 679

     226 ysSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
         ||||||||||||||:::|||||||||||||||||||||||||||||||||
     680 GTTCAGGTGTTTATGCAGGGGATCCTTCAAAATTAAGCATGAAAGCAGCA 729

     243 PheGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGl 259
         ||||||:::||||||||||||||||||   ||||||||||||||||||||
     730 TTTGGAAGAGTATGGAAGCTAGAAGAGATTGGTGGCAGCATCATTGGTGG 779

     259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
         ||||||||||:::|||:::||||||::::::|||||||||   |||||||
     780 CACTTTCAAGACAATCCAGGAGAGAAATAAGACACCTAAGCCACCCAGAG 829

     276 spProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 292
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     830 ACCCGCGTCTGCCAAAACCGAAGGGCCAAACAGTTGGATCTTTTAGGAAG 879

     293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
         ||||||   |||||||||:::||||||:::      |||||||||:::::
     880 GGACTTACCATGCTGCCTGAGGCAATTGCTAACAGTTTGGGTAGCAATGT 929

     309 uLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyT 326
         :|||||||||||||||||||||||||||||||||      :::|||||||
     930 AAAATTATCTTGGAAGCTTTCCAGTATTACCAAATTGGGCAATGGAGGGT 979

     326 yrHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
         ||:::||||||:::|||||||||||||||:::||||||||||||||||||
     980 ATAACTTGACATTTGAAACACCTGAAGGAATGGTATCTCTTCAGAGTAGA 1029

     343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
         |||:::|||||||||:::||||||:::||||||||||||:::|||:::||
    1030 AGTGTTGTAATGACCATTCCATCCCATGTTGCCAGTAACTTGTTGCATCC 1079

     359 oLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProV 376
         |||||||:::|||||||||||||||||||||:::||||||||||||||||
    1080 TCTCTCGGCTGCTGCTGCAGATGCATTATCCCAATTTTATTATCCTCCAG 1129

     376 alGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArg 392
         ||::::::||||||:::|||||||||:::||||||||||||   |||
    1130 TTGCATCAGTCACAGTCTCCTATCCAAAAGAAGCCATTCGAAAAGAATGT 1179

     393 LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGl 409
         |||:::|||||||||||||||||||||||||||||||||||||||:::||
    1180 TTGATTGATGGTGAACTTAAGGGGTTTGGCCAGTTGCACCCACGCAGCCA 1229

     409 nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnA 426
         ||||:::|||||||||||||||||||||||||||||||||||||||||||
    1230 AGGAATTGAAACTTTAGGGACGATATACAGTTCATCACTTTTCCCCAATC 1279

     426 rgAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
         ||||||||:::||||||||||||||||||||||||||||||||||||
    1280 GAGCTCCATCTGGCAGGGTGTTGCTCTTGAACTACATAGGAGGAGCTACC 1329

     443 AsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAs 459
         |||      ||||||||||||||||||::::::|||||||||:::|||||
    1330 AACACTGGAATTTTGTCCAAGACTGAAGGGGAACTTGTAGAAGCAGTTGA 1379

     459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
         |||||||||||||||||||||||||:::|||:::|||:::||||||||||
    1380 TCGTGATTTGAGAAAAATGCTTATAAATCCTAATGCAAAGGATCCTCTTG 1429

     476 alValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGly 492
         ||:::||||||||||||||||||:::||||||||||||||||||||||||
    1430 TTTTGGGTGTAAGAGTATGGCCAAAAGCCATTCCACAGTTCTTGGTTGGT 1479

     493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe 509
         |||||||||   |||::::::||||||   |||::::::|||:::|||::
    1480 CATTTGGATCTCCTTGATAGTGCAAAAATGGCTCTCAGGGATTCTGGGTT 1529

     509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
         ::::||||||||||||||||||||||||||||||||||||||||||||||
    1530 TCATGGACTGTTTCTTGGGGGCAACTATGTATCTGGTGTGGCATTAGGAC 1579

     526 rgCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 542
         |||||||||||||||||||||||||||||:::||||||      ||||||
    1580 GGTGTGTGGAAGGTGCTTACGAGGTTGCAGCTGAAGTGAAGGAATTCCTG 1629

     543 SerArgTyrAlaTyrLys 548
         |||:::||||||||||||
    1630 TCACAATATGCATACAAA 1647

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-15

seq_documentation_block:
; Sequence 15, Application US/08808323
; Patent No. 6018105
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Ward, Eric
;    TITLE OF INVENTION:  Promoters from Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase Genes
```

```
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6018105artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,323
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1846
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1826 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Gossypium hirsutum (cotton)
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-15 (NRRL B-21594)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  31..1647
;      OTHER INFORMATION:  /product= "Cotton protox-1 coding
;      OTHER INFORMATION:  region"
US-08-808-323-15


alignment_scores:
           Quality:  2113.00             Length:     506
             Ratio:  4.477                 Gaps:       1
 Percent Similarity:  93.281   Percent Identity:  79.249

alignment_block:
 US-09-508-418-2 x US-08-808-323-15    ..

 Align seg 1/1  to: US-08-808-323-15  from: 1  to: 1826

     47 TrpArgThrArgCysSerValAlaLysAspTyrThrValProSerSerAl 63
        ::::::   |||||||||:::|||:::      |||:::   ||||||
    130 TTCAAGCTCCGATGCTCCCTCGCCGAGGGTCCCACGATTTCCTCATCTAA 179

     63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
         :::|||||||||   ::::::      ||||||||||||||||||:::|
    180 AATCGACGGGGGAGAATCATCCATCGCGGATTGCGTCATCGTTGGAGGTG 229

     80 lyIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr...... 94
        ||||||||||||||||||||||||||::::::::::::::::::
    230 GTATCAGTGGACTTTGCATTGCTCAAGCTCTCGCCACCAAGCACCGTGAC 279

     95 ......ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAs 109
                 |||::::::|||||||||||||||||||||:::||||||||
    280 GTCGCTTCCAATGTGATTGTGACGGAGGCCAGAGACCGTGTTGGTGGCAA 329

    109 nIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnS 126
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    330 CATCACTACCGTTGAGAGAGATGGATATCTGTGGGAAGAAGGCCCCAACA 379

    126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
        ||||||||||||||||||||:::||||||||||||||||||   ||||||
    380 GTTTTCAGCCCTCCGATCCTATTCTAACCATGGCCGTGGATAGTGGATTG 429

    143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    430 AAGGACGATTTGGTTTTAGGTGACCCTAATGCACCGCGATTTGTACTATG 479

    159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
        |:::|||||||||||||||||||||||||||   ||||||||||||||||
    480 GGAGGGAAAACTAAGGCCTGTGCCCTCCAAGCCAACCGACTTGCCGTTTT 529

    176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
        |||||||||||||||||   ||||||||||||||||||||||||   |||
    530 TTGATTTGATGAGCATTGCTGGAAAACTTAGGGCTGGGTTCGGGGCTATT 579

    193 GlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVa 209
        |||:::||||||   |||||||||:::|||||||||||||||:::|||||
    580 GGCATTCGGCCTCCCCCTCCGGGTTATGAAGAATCGGTGGAGGAGTTTGT 629

    209 lArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC 226
        ||||||||||||||||:::|||||||||||||||:::|||||||||||||
    630 GCGCCGTAATCTTGGTGCTGAGGTTTTTGAACGCTTTATTGAACCATTTT 679

    226 ysSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
        ||||||||||||||:::|||||||||||||||||||||||||||||||||
    680 GTTCAGGTGTTTATGCAGGGGATCCTTCAAAATTAAGCATGAAAGCAGCA 729

    243 PheGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGl 259
        ||||||:::||||||||||||||||||   ||||||||||||||||||||
    730 TTTGGAAGAGTATGGAAGCTAGAAGAGATTGGTGGCAGCATCATTGGTGG 779

    259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
        ||||||||||:::|||:::||||||::::::|||||||||   |||||||
    780 CACTTTCAAGACAATCCAGGAGAGAAATAAGACACCTAAGCCACCCAGAG 829

    276 spProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 292
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    830 ACCCGCGTCTGCCAAAACCGAAGGGCCAAACAGTTGGATCTTTTAGGAAG 879

    293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
        ||||||   |||||||||:::||||||:::      |||||||||:::::
    880 GGACTTACCATGCTGCCTGAGGCAATTGCTAACAGTTTGGGTAGCAATGT 929

    309 uLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyT 326
        :|||||||||||||||||||||||||||||||||      :::|||||||
    930 AAAATTATCTTGGAAGCTTTCCAGTATTACCAAATTGGGCAATGGAGGGT 979

    326 yrHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
        ||:::||||||:::|||||||||||||||:::||||||||||||||||||
    980 ATAACTTGACATTTGAAACACCTGAAGGAATGGTATCTCTTCAGAGTAGA 1029

    343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
        |||:::|||||||||:::||||||:::||||||||||||:::|||:::||
   1030 AGTGTTGTAATGACCATTCCATCCCATGTTGCCAGTAACTTGTTGCATCC 1079

    359 oLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProV 376
        |||||||:::|||||||||||||||||||||:::||||||||||||||||
   1080 TCTCTCGGCTGCTGCTGCAGATGCATTATCCCAATTTTATTATCCTCCAG 1129

    376 alGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArg 392
        ||::::::||||||:::|||||||||:::||||||||||||   |||
```

```
1130 TTGCATCAGTCACAGTCTCCTATCCAAAAGAAGCCATTCGAAAAGAATGT 1179

 393 LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGl 409
     |||:::|||||||||||||||||||||||||||||||||||||||:::||
1180 TTGATTGATGGTGAACTTAAGGGGTTTGGCCAGTTGCACCCACGCAGCCA 1229

 409 nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnA 426
     ||||:::|||||||||||||||||||||||||||||||||||||||||||
1230 AGGAATTGAAACTTTAGGGACGATATACAGTTCATCACTTTTCCCCAATC 1279

 426 rgAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
     ||||||||:::||||||||||||||||||||||||||||||||||||
1280 GAGCTCCATCTGGCAGGGTGTTGCTCTTGAACTACATAGGAGGAGCTACC 1329

 443 AsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAs 459
     |||      ||||||||||||||||||::::::|||||||||:::|||||
1330 AACACTGGAATTTTGTCCAAGACTGAAGGGGAACTTGTAGAAGCAGTTGA 1379

 459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
     |||||||||||||||||||||||||:::|||:::|||:::||||||||||
1380 TCGTGATTTGAGAAAAATGCTTATAAATCCTAATGCAAAGGATCCTCTTG 1429

 476 alValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGly 492
     ||:::||||||||||||||||||:::||||||||||||||||||||||||
1430 TTTTGGGTGTAAGAGTATGGCCAAAAGCCATTCCACAGTTCTTGGTTGGT 1479

 493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe 509
     |||||||||   |||::::::||||||   |||::::::|||:::|||::
1480 CATTTGGATCTCCTTGATAGTGCAAAAATGGCTCTCAGGGATTCTGGGTT 1529

 509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
     ::::||||||||||||||||||||||||||||||||||||||||||||||
1530 TCATGGACTGTTTCTTGGGGGCAACTATGTATCTGGTGTGGCATTAGGAC 1579

 526 rgCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 542
     |||||||||||||||||||||||||||||:::||||||      ||||||
1580 GGTGTGTGGAAGGTGCTTACGAGGTTGCAGCTGAAGTGAAGGAATTCCTG 1629

 543 SerArgTyrAlaTyrLys 548
     |||:::||||||||||||
1630 TCACAATATGCATACAAA 1647

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-15

seq_documentation_block:
; Sequence 15, Application US/09050603A
; Patent No. 6023012
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6023012artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/050,603A
;      FILING DATE:  30-MAR-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1826 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Gossypium hirsutum (cotton)
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-15 (NRRL B-21594)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  31..1647
;      OTHER INFORMATION:  /product= "Cotton protox-1 coding
;      OTHER INFORMATION:  sequence"
US-09-050-603A-15


alignment_scores:
            Quality: 2113.00              Length:    506
              Ratio:  4.477                 Gaps:      1
 Percent Similarity:  93.281   Percent Identity: 79.249

alignment_block:
 US-09-508-418-2 x US-09-050-603A-15    ..

 Align seg 1/1  to: US-09-050-603A-15  from: 1  to: 1826

  47 TrpArgThrArgCysSerValAlaLysAspTyrThrValProSerSerAl 63
     ::::::   |||||||||:::|||:::      |||:::   ||||||
 130 TTCAAGCTCCGATGCTCCCTCGCCGAGGGTCCCACGATTTCCTCATCTAA 179

  63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
      :::|||||||||   ::::::      ||||||||||||||||||:::|
 180 AATCGACGGGGGAGAATCATCCATCGCGGATTGCGTCATCGTTGGAGGTG 229

  80 lyIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr...... 94
     ||||||||||||||||||||||||||::::::::::::::::::
 230 GTATCAGTGGACTTTGCATTGCTCAAGCTCTCGCCACCAAGCACCGTGAC 279

  95 ......ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAs 109
              |||::::::|||||||||||||||||||||:::||||||||
 280 GTCGCTTCCAATGTGATTGTGACGGAGGCCAGAGACCGTGTTGGTGGCAA 329

 109 nIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnS 126
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 330 CATCACTACCGTTGAGAGAGATGGATATCTGTGGGAAGAAGGCCCCAACA 379

 126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
     ||||||||||||||||||||:::||||||||||||||||||   ||||||
 380 GTTTTCAGCCCTCCGATCCTATTCTAACCATGGCCGTGGATAGTGGATTG 429
```

```
 143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 430 AAGGACGATTTGGTTTTAGGTGACCCTAATGCACCGCGATTTGTACTATG 479

 159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
     |:::|||||||||||||||||||||||||||   ||||||||||||||||
 480 GGAGGGAAAACTAAGGCCTGTGCCCTCCAAGCCAACCGACTTGCCGTTTT 529

 176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
     |||||||||||||||||   ||||||||||||||||||||||||   |||
 530 TTGATTTGATGAGCATTGCTGGAAAACTTAGGGCTGGGTTCGGGGCTATT 579

 193 GlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVa 209
     |||:::||||||   |||||||||:::|||||||||||||||:::|||||
 580 GGCATTCGGCCTCCCCCTCCGGGTTATGAAGAATCGGTGGAGGAGTTTGT 629

 209 lArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC 226
     ||||||||||||||||:::|||||||||||||||:::|||||||||||||
 630 GCGCCGTAATCTTGGTGCTGAGGTTTTTGAACGCTTTATTGAACCATTTT 679

 226 ysSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
     ||||||||||||||:::|||||||||||||||||||||||||||||||||
 680 GTTCAGGTGTTTATGCAGGGGATCCTTCAAAATTAAGCATGAAAGCAGCA 729

 243 PheGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGl 259
     ||||||:::||||||||||||||||||   ||||||||||||||||||||
 730 TTTGGAAGAGTATGGAAGCTAGAAGAGATTGGTGGCAGCATCATTGGTGG 779

 259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
     ||||||||||:::|||:::||||||::::::|||||||||   |||||||
 780 CACTTTCAAGACAATCCAGGAGAGAAATAAGACACCTAAGCCACCCAGAG 829

 276 spProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 292
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 830 ACCCGCGTCTGCCAAAACCGAAGGGCCAAACAGTTGGATCTTTTAGGAAG 879

 293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
     ||||||   |||||||||:::||||||:::      |||||||||:::::
 880 GGACTTACCATGCTGCCTGAGGCAATTGCTAACAGTTTGGGTAGCAATGT 929

 309 uLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyT 326
     :|||||||||||||||||||||||||||||||||      :::|||||||
 930 AAAATTATCTTGGAAGCTTTCCAGTATTACCAAATTGGGCAATGGAGGGT 979

 326 yrHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
     ||:::||||||:::|||||||||||||||:::||||||||||||||||||
 980 ATAACTTGACATTTGAAACACCTGAAGGAATGGTATCTCTTCAGAGTAGA 1029

 343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
     |||:::|||||||||:::||||||:::||||||||||||:::|||:::||
1030 AGTGTTGTAATGACCATTCCATCCCATGTTGCCAGTAACTTGTTGCATCC 1079

 359 oLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProV 376
     |||||||:::|||||||||||||||||||||:::||||||||||||||||
1080 TCTCTCGGCTGCTGCTGCAGATGCATTATCCCAATTTTATTATCCTCCAG 1129

 376 alGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArg 392
     ||::::::||||||:::|||||||||:::||||||||||||   |||
1130 TTGCATCAGTCACAGTCTCCTATCCAAAAGAAGCCATTCGAAAAGAATGT 1179

 393 LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGl 409
     |||:::|||||||||||||||||||||||||||||||||||||||:::||
1180 TTGATTGATGGTGAACTTAAGGGGTTTGGCCAGTTGCACCCACGCAGCCA 1229

 409 nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnA 426
     ||||:::|||||||||||||||||||||||||||||||||||||||||||
1230 AGGAATTGAAACTTTAGGGACGATATACAGTTCATCACTTTTCCCCAATC 1279

 426 rgAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
     ||||||||:::||||||||||||||||||||||||||||||||||||
1280 GAGCTCCATCTGGCAGGGTGTTGCTCTTGAACTACATAGGAGGAGCTACC 1329

 443 AsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAs 459
     |||      ||||||||||||||||||::::::|||||||||:::|||||
1330 AACACTGGAATTTTGTCCAAGACTGAAGGGGAACTTGTAGAAGCAGTTGA 1379

 459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
     |||||||||||||||||||||||||:::|||:::|||:::||||||||||
1380 TCGTGATTTGAGAAAAATGCTTATAAATCCTAATGCAAAGGATCCTCTTG 1429

 476 alValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGly 492
     ||:::||||||||||||||||||:::||||||||||||||||||||||||
1430 TTTTGGGTGTAAGAGTATGGCCAAAAGCCATTCCACAGTTCTTGGTTGGT 1479

 493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe 509
     |||||||||   |||::::::||||||   |||::::::|||:::|||::
1480 CATTTGGATCTCCTTGATAGTGCAAAAATGGCTCTCAGGGATTCTGGGTT 1529

 509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
     ::::||||||||||||||||||||||||||||||||||||||||||||||
1530 TCATGGACTGTTTCTTGGGGGCAACTATGTATCTGGTGTGGCATTAGGAC 1579

 526 rgCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 542
     |||||||||||||||||||||||||||||:::||||||      ||||||
1580 GGTGTGTGGAAGGTGCTTACGAGGTTGCAGCTGAAGTGAAGGAATTCCTG 1629

 543 SerArgTyrAlaTyrLys 548
     |||:::||||||||||||
1630 TCACAATATGCATACAAA 1647
```

```
seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-15

seq_documentation_block:
; Sequence 15, Application US/09102420B
; Patent No. 6084155
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
;    TITLE OF INVENTION:  OXIDASE ("PROTOX")
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6084155artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/102,420B
;      FILING DATE:  22-JUN-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/059,164
;      FILING DATE:  13-APR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/050,603
;      FILING DATE:  30-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/126,430
;      FILING DATE:  11-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
```

```
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/472,028
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847/CIP3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1826 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Gossypium hirsutum (cotton)
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-15 (NRRL B-21594)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  31..1647
;      OTHER INFORMATION:  /product= "Cotton protox-1 coding
;      OTHER INFORMATION:  sequence"
US-09-102-420B-15


alignment_scores:
            Quality: 2113.00            Length:    506
              Ratio:  4.477               Gaps:      1
 Percent Similarity:  93.281   Percent Identity: 79.249

alignment_block:
 US-09-508-418-2 x US-09-102-420B-15   ..

 Align seg 1/1  to: US-09-102-420B-15  from: 1  to: 1826

      47 TrpArgThrArgCysSerValAlaLysAspTyrThrValProSerSerAl 63
         ::::::   |||||||||:::|||:::      |||:::   ||||||
     130 TTCAAGCTCCGATGCTCCCTCGCCGAGGGTCCCACGATTTCCTCATCTAA 179

      63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
          :::|||||||||   ::::::      ||||||||||||||||||:::|
     180 AATCGACGGGGGAGAATCATCCATCGCGGATTGCGTCATCGTTGGAGGTG 229

      80 lyIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr...... 94
         ||||||||||||||||||||||||||::::::::::::::::::
     230 GTATCAGTGGACTTTGCATTGCTCAAGCTCTCGCCACCAAGCACCGTGAC 279

      95 ......ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAs 109
                  |||::::::|||||||||||||||||||||:::||||||||
     280 GTCGCTTCCAATGTGATTGTGACGGAGGCCAGAGACCGTGTTGGTGGCAA 329

     109 nIleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnS 126
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     330 CATCACTACCGTTGAGAGAGATGGATATCTGTGGGAAGAAGGCCCCAACA 379

     126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
         ||||||||||||||||||||:::||||||||||||||||||   ||||||
     380 GTTTTCAGCCCTCCGATCCTATTCTAACCATGGCCGTGGATAGTGGATTG 429

     143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     430 AAGGACGATTTGGTTTTAGGTGACCCTAATGCACCGCGATTTGTACTATG 479
```

```
     159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
         |:::|||||||||||||||||||||||||||   ||||||||||||||||
     480 GGAGGGAAAACTAAGGCCTGTGCCCTCCAAGCCAACCGACTTGCCGTTTT 529

     176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
         |||||||||||||||||   ||||||||||||||||||||||||   |||
     530 TTGATTTGATGAGCATTGCTGGAAAACTTAGGGCTGGGTTCGGGGCTATT 579

     193 GlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVa 209
         |||:::||||||   |||||||||:::|||||||||||||||:::|||||
     580 GGCATTCGGCCTCCCCCTCCGGGTTATGAAGAATCGGTGGAGGAGTTTGT 629

     209 lArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC 226
         ||||||||||||||||:::|||||||||||||||:::|||||||||||||
     630 GCGCCGTAATCTTGGTGCTGAGGTTTTTGAACGCTTTATTGAACCATTTT 679

     226 ysSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
         ||||||||||||||:::|||||||||||||||||||||||||||||||||
     680 GTTCAGGTGTTTATGCAGGGGATCCTTCAAAATTAAGCATGAAAGCAGCA 729

     243 PheGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGl 259
         ||||||:::||||||||||||||||||   ||||||||||||||||||||
     730 TTTGGAAGAGTATGGAAGCTAGAAGAGATTGGTGGCAGCATCATTGGTGG 779

     259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
         ||||||||||:::|||:::||||||::::::|||||||||   |||||||
     780 CACTTTCAAGACAATCCAGGAGAGAAATAAGACACCTAAGCCACCCAGAG 829

     276 spProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 292
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     830 ACCCGCGTCTGCCAAAACCGAAGGGCCAAACAGTTGGATCTTTTAGGAAG 879

     293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
         ||||||   |||||||||:::||||||:::      |||||||||:::::
     880 GGACTTACCATGCTGCCTGAGGCAATTGCTAACAGTTTGGGTAGCAATGT 929

     309 uLysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyT 326
         :|||||||||||||||||||||||||||||||||      :::|||||||
     930 AAAATTATCTTGGAAGCTTTCCAGTATTACCAAATTGGGCAATGGAGGGT 979

     326 yrHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
         ||:::||||||:::|||||||||||||||:::||||||||||||||||||
     980 ATAACTTGACATTTGAAACACCTGAAGGAATGGTATCTCTTCAGAGTAGA 1029

     343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
         |||:::|||||||||:::||||||:::||||||||||||:::|||:::||
    1030 AGTGTTGTAATGACCATTCCATCCCATGTTGCCAGTAACTTGTTGCATCC 1079

     359 oLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProV 376
         |||||||:::|||||||||||||||||||||:::||||||||||||||||
    1080 TCTCTCGGCTGCTGCTGCAGATGCATTATCCCAATTTTATTATCCTCCAG 1129

     376 alGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArg 392
         ||::::::||||||:::|||||||||:::||||||||||||   |||
    1130 TTGCATCAGTCACAGTCTCCTATCCAAAAGAAGCCATTCGAAAAGAATGT 1179

     393 LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGl 409
         |||:::|||||||||||||||||||||||||||||||||||||||:::||
    1180 TTGATTGATGGTGAACTTAAGGGGTTTGGCCAGTTGCACCCACGCAGCCA 1229

     409 nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnA 426
         ||||:::|||||||||||||||||||||||||||||||||||||||||||
    1230 AGGAATTGAAACTTTAGGGACGATATACAGTTCATCACTTTTCCCCAATC 1279

     426 rgAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
         ||||||||:::||||||||||||||||||||||||||||||||||||
    1280 GAGCTCCATCTGGCAGGGTGTTGCTCTTGAACTACATAGGAGGAGCTACC 1329

     443 AsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAs 459
         |||      ||||||||||||||||||::::::|||||||||:::|||||
    1330 AACACTGGAATTTTGTCCAAGACTGAAGGGGAACTTGTAGAAGCAGTTGA 1379

     459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
```

```
     ||||||||||||||||||||||||||:::|||:::|||:::|||||||||
1380 TCGTGATTTGAGAAAAATGCTTATAAATCCTAATGCAAAGGATCCTCTTG 1429

 476 alValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGly 492
     ||:::||||||||||||||||||:::||||||||||||||||||||||||
1430 TTTTGGGTGTAAGAGTATGGCCAAAAGCCATTCCACAGTTCTTGGTTGGT 1479

 493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe 509
     |||||||||   |||::::::||||||   |||::::::|||:::|||::
1480 CATTTGGATCTCCTTGATAGTGCAAAAATGGCTCTCAGGGATTCTGGGTT 1529

 509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
     ::::||||||||||||||||||||||||||||||||||||||||||||||
1530 TCATGGACTGTTTCTTGGGGGCAACTATGTATCTGGTGTGGCATTAGGAC 1579

 526 rgCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 542
     |||||||||||||||||||||||||||||:::||||||      ||||||
1580 GGTGTGTGGAAGGTGCTTACGAGGTTGCAGCTGAAGTGAAGGAATTCCTG 1629

 543 SerArgTyrAlaTyrLys 548
     |||:::||||||||||||
1630 TCACAATATGCATACAAA 1647

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-472-028A-1

seq_documentation_block:
; Sequence 1, Application US/08472028A
; Patent No. 5767373
;  GENERAL INFORMATION:
;    APPLICANT:  Ward, Eric R
;    APPLICANT:  Volrath, Sandra
;    TITLE OF INVENTION:  Manipulation of Protoporphyrinogen
;    TITLE OF INVENTION:  Oxidase Enzyme Activity in Eukaryotic Organisms
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ciba-Geigy Corporation
;      STREET:  7 Skyline Drive
;      CITY:  Hawthorne
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10532
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/472,028A
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Elmer, James Scott
;      REGISTRATION NUMBER:  36,129
;      REFERENCE/DOCKET NUMBER:  CGC 1748/CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8614
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1719 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  31..1644
;      OTHER INFORMATION:  /note= "Arabidopsis protox-1 cDNA;
;      OTHER INFORMATION:  sequence from pWDC-2"
US-08-472-028A-1
```

```
alignment_scores:
          Quality:  2077.00              Length:    539
            Ratio:  4.230                  Gaps:      4
 Percent Similarity:  91.095   Percent Identity:  75.139

alignment_block:
 US-09-508-418-2 x US-08-472-028A-1   ..

 Align seg 1/1  to: US-08-472-028A-1  from: 1  to: 1719

  19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
     |||      |||::::::|||:::   |||:::         :::   ||
  28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

  35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
     |:::||||||   |||   ::::::   |||   :::         |||
  78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

  49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
       |||||||||||||||         ||||||   ||||||   ::::::
 125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

  66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
     ||||||   ::::::   :::      ||||||||||||||||||:::||
 175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

  80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
     |||||||||||||||||||||||||::::::::::::::::::
 225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274

  95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
          |||||||||:::||||||||||||:::||||||:::|||||||||
 275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324

 110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
     |||   |||   |||::::::|||:::|||||||||||||||||||||||
 325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

 126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
     |||||||||||||||||||||||||||||||:::||||||   |||||||
 375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

 143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
     ||||||||||||||||||||||||||:::|||||||||||||||||||||
 425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

 160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
     :::|||||||||||||||||||||||||||||||||||||||||||||||
 475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524

 176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
     ||||||||||||||||   ||||||:::|||||||||||||||   :::|
 525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

 193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
     ||:::||||||||||||||||||:::|||||||||||||||:::||||||
 575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

 210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
     |||||||||||||||   ||||||||||||||||||||||||||||||||
 625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674

 226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
     |||||||||||||:::||||||||||||||||||||||||||||||||||
 675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

 243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
     |||||||||||||||||||||||::::::|||||||||||||||||||||
 725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

 260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
     |||||||||||||||:::||||||:::::::::|||||||||   |||||
```

```
  775 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

  276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
      |||||||||||||||||||:::||||||||||||||||||||||||||||
  825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 874

  293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
      |||||||||||||||||:::|||||||||||||||||||||||||||:::
  875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

  310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
      |||||||||||||||||||||:::|||||||||   |||:::||||||||
  925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

  326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
      |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
  975 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1024

  343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
      ||:::||||||||||||||||||:::|||||||||::::::|||||||||
 1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

  360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
      ||||||   :::||||||:::|||||||||::::::||||||||||||||
 1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

  376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
      |:::||||||:::||||||||||||:::||||||||||||   |||   |
 1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

  393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
      ||:::|||||||||||||||||||||||||||||||||||||||||||||
 1175 TGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAA 1224

  410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274

  426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
      |||||||   ||||||:::||||||||||||||||||||||||:::   |
 1275 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

  443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
      ||      ||||||||||||:::|||::::::|||||||||:::||||||
 1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

  460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
      ||||||||||||||||||||||||||||||::::::   |||||||||
 1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

  476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
       :::||||||||||||||||||||||||||||||||||||||||||||||
 1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

  493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
      ||:::|||   |||:::|||||||||::::::::::::::::::|||
 1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

  510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
      ||||||||||||||||||||||||||||||:::|||||||||||||||||
 1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

  526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
      ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
 1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

  543 erArgTyrAlaTyrLys 548
      |||||||||||||||||
 1625 CACGGTACGCTTACAAG 1641

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-1

seq_documentation_block:
```

```
; Sequence 1, Application US/08808931
; Patent No. 5939602
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5939602artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,931
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1719 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Arabidopsis thaliana
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-2 (NRRL B-21238)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  31..1644
;      OTHER INFORMATION:  /product= "Arabidopsis protox-1"
US-08-808-931-1


alignment_scores:
           Quality: 2077.00            Length:    539
             Ratio:  4.230               Gaps:      4
 Percent Similarity:  91.095  Percent Identity: 75.139

alignment_block:
 US-09-508-418-2 x US-08-808-931-1   ..

 Align seg 1/1  to: US-08-808-931-1  from: 1  to: 1719
```

```
 19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
    |||      |||::::::|||:::   |||:::         :::   ||
 28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

 35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
    |:::||||||   |||   ::::::   |||   :::         |||
 78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

 49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
      |||||||||||||||         ||||||   ||||||   ::::::
125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

 66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
    ||||||   ::::::   :::      ||||||||||||||||||:::||
175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

 80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
    |||||||||||||||||||||||||::::::::::::::::::
225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274

 95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
         |||||||||:::||||||||||||:::||||||:::|||||||||
275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324

110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
    |||   |||   |||::::::|||:::|||||||||||||||||||||||
325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
    |||||||||||||||||||||||||||||||:::||||||   |||||||
375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
    ||||||||||||||||||||||||||:::|||||||||||||||||||||
425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
    :::|||||||||||||||||||||||||||||||||||||||||||||||
475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524

176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
    ||||||||||||||||   ||||||:::|||||||||||||||   :::|
525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
    ||:::||||||||||||||||||:::|||||||||||||||:::||||||
575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
    |||||||||||||||   ||||||||||||||||||||||||||||||||
625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674

226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
    |||||||||||||:::||||||||||||||||||||||||||||||||||
675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
    |||||||||||||||||||||||::::::|||||||||||||||||||||
725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
    |||||||||||||||:::||||||:::::::::|||||||||   |||||
775 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
    |||||||||||||||||||:::||||||||||||||||||||||||||||
825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 874

293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
    |||||||||||||||||:::|||||||||||||||||||||||||||:::
875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

 310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
     |||||||||||||||||||||:::|||||||||   |||:::||||||||
 925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

 326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
     |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
 975 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1024

 343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
     ||:::||||||||||||||||||:::|||||||||::::::|||||||||
1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
     ||||||   :::||||||:::|||||||||::::::||||||||||||||
1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

 376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
     |:::||||||:::||||||||||||:::||||||||||||   |||   |
1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

 393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
     ||:::|||||||||||||||||||||||||||||||||||||||||||||
1175 TGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAA 1224

 410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274

 426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
     |||||||   ||||||:::||||||||||||||||||||||||:::   |
1275 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

 443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
     ||      ||||||||||||:::|||::::::|||||||||:::||||||
1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

 460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
     ||||||||||||||||||||||||||||||::::::   |||||||||
1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

 476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
      :::||||||||||||||||||||||||||||||||||||||||||||||
1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

 493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
     ||:::|||   |||:::|||||||||::::::::::::::::::|||
1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

 510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
     ||||||||||||||||||||||||||||||:::|||||||||||||||||
1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

 526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
     ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

 543 erArgTyrAlaTyrLys 548
     |||||||||||||||||
1625 CACGGTACGCTTACAAG 1641

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-1

seq_documentation_block:
; Sequence 1, Application US/08808323
; Patent No. 6018105
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Ward, Eric
;    TITLE OF INVENTION:  Promoters from Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase Genes
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  No. 6018105artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,323
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1846
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1719 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Arabidopsis thaliana
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-2 (NRRL B-21238)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  31..1644
;      OTHER INFORMATION:  /product= "Arabidopsis protox-1"
US-08-808-323-1

alignment_scores:
          Quality:  2077.00          Length:     539
            Ratio:     4.230           Gaps:       4
 Percent Similarity:  91.095   Percent Identity:  75.139

alignment_block:
 US-09-508-418-2 x US-08-808-323-1    ..

 Align seg 1/1  to: US-08-808-323-1  from: 1  to: 1719

      19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
         |||      |||::::::|||:::   |||:::         :::   ||
      28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

      35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
         |:::||||||   |||   ::::::   |||   :::         |||
      78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

      49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
           |||||||||||||||         ||||||   ||||||   ::::::
   .  125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

  66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
     ||||||   ::::::   :::      ||||||||||||||||||:::||
 175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

  80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
     |||||||||||||||||||||||||::::::::::::::::::
 225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274

  95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
          |||||||||:::||||||||||||:::||||||:::|||||||||
 275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324

 110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
     |||   |||   |||::::::|||:::|||||||||||||||||||||||
 325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

 126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
     |||||||||||||||||||||||||||||||:::||||||   |||||||
 375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

 143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
     ||||||||||||||||||||||||||:::|||||||||||||||||||||
 425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

 160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
     :::|||||||||||||||||||||||||||||||||||||||||||||||
 475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524

 176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
     ||||||||||||||||   ||||||:::|||||||||||||||   :::|
 525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

 193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
     ||:::||||||||||||||||||:::|||||||||||||||:::||||||
 575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

 210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
     |||||||||||||||   ||||||||||||||||||||||||||||||||
 625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674

 226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
     |||||||||||||:::||||||||||||||||||||||||||||||||||
 675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

 243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
     |||||||||||||||||||||||::::::|||||||||||||||||||||
 725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

 260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
     |||||||||||||||:::||||||:::::::::|||||||||   |||||
 775 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

 276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
     |||||||||||||||||||:::||||||||||||||||||||||||||||
 825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 874

 293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
     |||||||||||||||||:::|||||||||||||||||||||||||||:::
 875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

 310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
     |||||||||||||||||||||:::|||||||||   |||:::||||||||
 925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

 326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
     |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
 975 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1024

 343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
     ||:::||||||||||||||||||:::|||||||||::::::|||||||||
1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
```

```
          ||||||   :::||||||:::|||||||||::::::||||||||||||||
     1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

      376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
          |:::||||||:::|||||||||||||||:::|||||||||   |||   |
     1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

      393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
          ||:::|||||||||||||||||||||||||||||||||||||||||||||
     1175 TGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAA 1224

      410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274

      426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
          |||||||   ||||||:::||||||||||||||||||||||||:::   |
     1275 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

      443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
          ||      ||||||||||||:::|||::::::|||||||||:::||||||
     1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

      460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
          ||||||||||||||||||||||||||||||::::::   |||||||||
     1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

      476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
           :::||||||||||||||||||||||||||||||||||||||||||||||
     1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

      493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
          ||:::|||   |||:::|||||||||::::::::::::::::::|||
     1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

      510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
          ||||||||||||||||||||||||||||||:::|||||||||||||||||
     1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

      526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
          ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
     1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

      543 erArgTyrAlaTyrLys 548
          |||||||||||||||||
     1625 CACGGTACGCTTACAAG 1641

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-1

seq_documentation_block:
; Sequence 1, Application US/09050603A
; Patent No. 6023012
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6023012artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/050,603A
;      FILING DATE:  30-MAR-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1719 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Arabidopsis thaliana
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-2 (NRRL B-21238)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  31..1644
;      OTHER INFORMATION:  /product= "Arabidopsis protox-1"
US-09-050-603A-1


alignment_scores:
            Quality: 2077.00             Length:    539
              Ratio:  4.230                Gaps:      4
 Percent Similarity:  91.095   Percent Identity: 75.139

alignment_block:
 US-09-508-418-2 x US-09-050-603A-1   ..

 Align seg 1/1  to: US-09-050-603A-1  from: 1  to: 1719

       19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
          |||      |||::::::|||:::   |||:::         :::   ||
       28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

       35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
          |:::||||||   |||   ::::::   |||   :::         |||
       78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

       49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
            |||||||||||||||         ||||||   ||||||   ::::::
      125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

       66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
          ||||||   ::::::   :::      ||||||||||||||||||:::||
      175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

       80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
          |||||||||||||||||||||||||::::::::::::::::::
      225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274
```

```
  95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
          |||||||||:::||||||||||||:::||||||:::|||||||||
 275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324

 110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
     |||   |||   |||::::::|||:::|||||||||||||||||||||||
 325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

 126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
     |||||||||||||||||||||||||||||||:::||||||   |||||||
 375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

 143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
     ||||||||||||||||||||||||||:::|||||||||||||||||||||
 425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

 160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
     :::|||||||||||||||||||||||||||||||||||||||||||||||
 475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524

 176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
     ||||||||||||||||   ||||||:::|||||||||||||||   :::|
 525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

 193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
     ||:::|||||||||||||||||||||:::||||||||||||:::||||||
 575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

 210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
     |||||||||||||||   ||||||||||||||||||||||||||||||||
 625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674

 226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
     |||||||||||||:::||||||||||||||||||||||||||||||||||
 675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

 243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
     |||||||||||||||||||||||::::::|||||||||||||||||||||
 725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

 260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
     |||||||||||||||:::||||||:::::::::|||||||||   |||||
 775 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

 276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
     ||||||||||||||||||||||:::|||||||||||||||||||||||||
 825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 874

 293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
     |||||||||||||||||:::|||||||||||||||||||||||||||:::
 875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

 310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
     |||||||||||||||||||||:::|||||||||   |||:::||||||||
 925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

 326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
     |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
 975 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1024

 343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
     ||:::||||||||||||||||||:::|||||||||::::::|||||||||
1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
     ||||||   :::||||||:::|||||||||::::::||||||||||||||
1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

 376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
     |:::||||||:::||||||||||||:::||||||||||||   |||   |
1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

 393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
```

```
     ||:::|||||||||||||||||||||||||||||||||||||||||||||
1175 TGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAA 1224

 410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274

 426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
     |||||||   ||||||:::||||||||||||||||||||||||:::   |
1275 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

 443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
     ||      ||||||||||||:::|||::::::|||||||||:::||||||
1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

 460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
     ||||||||||||||||||||||||||||||::::::   |||||||||
1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

 476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
      :::||||||||||||||||||||||||||||||||||||||||||||||
1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

 493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
     ||:::|||   |||:::|||||||||::::::::::::::::::|||
1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

 510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
     ||||||||||||||||||||||||||||||:::|||||||||||||||||
1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

 526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
     ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

 543 erArgTyrAlaTyrLys 548
     |||||||||||||||||
1625 CACGGTACGCTTACAAG 1641
```

```
seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-1

seq_documentation_block:
; Sequence 1, Application US/09102420B
; Patent No. 6084155
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
;    TITLE OF INVENTION:  OXIDASE ("PROTOX")
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6084155artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/102,420B
;      FILING DATE:  22-JUN-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/059,164
;      FILING DATE:  13-APR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/050,603
```

```
;      FILING DATE:  30-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/126,430
;      FILING DATE:  11-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/472,028
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847/CIP3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1719 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Arabidopsis thaliana
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-2 (NRRL B-21238)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  31..1644
;      OTHER INFORMATION:  /product= "Arabidopsis protox-1"
US-09-102-420B-1


alignment_scores:
           Quality: 2077.00           Length:    539
             Ratio:  4.230              Gaps:      4
 Percent Similarity:  91.095   Percent Identity: 75.139

alignment_block:
 US-09-508-418-2 x US-09-102-420B-1    ..

 Align seg 1/1  to: US-09-102-420B-1  from: 1  to: 1719

  19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
     |||      |||::::::|||:::   |||:::         :::   ||
  28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

  35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
     |:::||||||   |||   ::::::   |||   :::         |||
  78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

  49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
       |||||||||||||||         ||||||   ||||||   ::::::
 125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

  66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
     ||||||   ::::::   :::      ||||||||||||||||||:::||
 175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

  80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
     |||||||||||||||||||||||||::::::::::::::::::
 225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274

  95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
          |||||||||:::||||||||||||:::||||||:::|||||||||
 275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324

 110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
     |||   |||   |||::::::|||:::|||||||||||||||||||||||
 325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

 126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
     |||||||||||||||||||||||||||||||:::||||||   |||||||
 375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

 143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
     ||||||||||||||||||||||||||:::|||||||||||||||||||||
 425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

 160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
     :::|||||||||||||||||||||||||||||||||||||||||||||||
 475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524

 176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
     ||||||||||||||||   ||||||:::|||||||||||||||   :::|
 525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

 193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
     ||:::||||||||||||||||||:::|||||||||||||||:::||||||
 575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

 210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
     |||||||||||||||   ||||||||||||||||||||||||||||||||
 625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674

 226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
     |||||||||||||:::||||||||||||||||||||||||||||||||||
 675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

 243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
     |||||||||||||||||||||||::::::|||||||||||||||||||||
 725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

 260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
     |||||||||||||||:::||||||:::::::::|||||||||   |||||
 775 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

 276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
     |||||||||||||||||||:::||||||||||||||||||||||||||||
 825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 874

 293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
     |||||||||||||||||:::|||||||||||||||||||||||||||:::
 875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

 310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
     |||||||||||||||||||||:::|||||||||   |||:::||||||||
 925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

 326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
     |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
 975 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1024

 343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
     ||:::||||||||||||||||||:::|||||||||::::::|||||||||
1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
     ||||||   :::||||||:::|||||||||::::::||||||||||||||
1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

 376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
     |:::||||||:::||||||||||||:::||||||||||||   |||   |
```

```
 1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

  393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
      ||:::|||||||||||||||||||||||||||||||||||||||||||||
 1175 TGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAA 1224

  410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274

  426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
      |||||||   ||||||:::||||||||||||||||||||||||:::   |
 1275 CGCACCGCCCGGAAGAATTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

  443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
      ||      ||||||||||||:::|||::::::|||||||||:::||||||
 1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

  460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
      ||||||||||||||||||||||||||||||::::::   |||||||||
 1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

  476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
       :::||||||||||||||||||||||||||||||||||||||||||||||
 1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

  493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
      ||:::|||   |||:::|||||||||::::::::::::::::::|||
 1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

  510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
      ||||||||||||||||||||||||||||||:::|||||||||||||||||
 1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

  526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
      ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
 1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

  543 erArgTyrAlaTyrLys 548
      |||||||||||||||||
 1625 CACGGTACGCTTACAAG 1641

seq_name: /cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-071-296-1

seq_documentation_block:
; Sequence 1, Application US/09071296
; Patent No. 6177245
;  GENERAL INFORMATION:
;    APPLICANT:  Ward, Eric R
;    APPLICANT:  Volrath, Sandra
;    TITLE OF INVENTION:  Manipulation of Protoporphyrinogen
;    TITLE OF INVENTION:  Oxidase Enzyme Activity in Eukaryotic Organisms
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ciba-Geigy Corporation
;      STREET:  7 Skyline Drive
;      CITY:  Hawthorne
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10532
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,296
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/261,198
;      FILING DATE:  16-JUN-94
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Elmer, James Scott
;      REGISTRATION NUMBER:  36,129
;      REFERENCE/DOCKET NUMBER:  CGC 1748/CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8614
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1719 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  31..1644
;      OTHER INFORMATION:  /note= "Arabidopsis protox-1 cDNA;
;      OTHER INFORMATION:  sequence from pWDC-2"
US-09-071-296-1


alignment_scores:
            Quality: 2077.00            Length:    539
              Ratio:  4.230               Gaps:      4
 Percent Similarity:  91.095   Percent Identity: 75.139

alignment_block:
 US-09-508-418-2 x US-09-071-296-1   ..

 Align seg 1/1  to: US-09-071-296-1  from: 1  to: 1719

   19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
      |||      |||::::::|||:::   |||:::         :::   ||
   28 TCCATGGAGTTATCTCTTCTCCGTCCGACGACTCAATCGCTTCTTCCGTC 77

   35 rIleSerLysArgAsnSerValAsnCysAsnGlyTrp.........ArgT 49
      |:::||||||   |||   ::::::   |||   :::         |||
   78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAAGCCTCTTAGAC 124

   49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
        |||||||||||||||         ||||||   ||||||   ::::::
  125 TCCGTTGTTCAGTGGCCGGTGGACCAACCGTCGGATCTTCAAAAATCGAA 174

   66 GlyGlyProAlaAlaGluLeu......AspCysValIleValGlyAlaGl 80
      ||||||   ::::::   :::      ||||||||||||||||||:::||
  175 GGCGGAGGAGGCACCACCATCACGACGGATTGTGTGATTGTCGGCGGAGG 224

   80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr....... 94
      |||||||||||||||||||||||||::::::::::::::::::
  225 TATTAGTGGTCTTTGCATCGCTCAGGCGCTTGCTACTAAGCATCCTGATG 274

   95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
           |||||||||:::||||||||||||:::||||||:::|||||||||
  275 CTGCTCCGAATTTAATTGTGACCGAGGCTAAGGATCGTGTTGGAGGCAAC 324

  110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
      |||   |||   |||::::::|||:::|||||||||||||||||||||||
  325 ATTATCACTCGTGAAGAGAATGGTTTTCTCTGGGAAGAAGGTCCCAATAG 374

  126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
      |||||||||||||||||||||||||||||||:::||||||   |||||||
  375 TTTTCAACCGTCTGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGA 424

  143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
      ||||||||||||||||||||||||||:::|||||||||||||||||||||
  425 AGGATGATTTGGTGTTGGGAGATCCTACTGCGCCAAGGTTTGTGTTGTGG 474

  160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
      :::|||||||||||||||||||||||||||||||||||||||||||||||
  475 AATGGGAAATTGAGGCCGGTTCCATCGAAGCTAACAGACTTACCGTTCTT 524
```

```
 176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
     ||||||||||||||||   ||||||:::|||||||||||||||   :::|
 525 TGATTTGATGAGTATTGGTGGGAAGATTAGAGCTGGTTTTGGTGCACTTG 574

 193 lyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
     ||:::||||||||||||||||||:::|||||||||||||||:::||||||
 575 GCATTCGACCGTCACCTCCAGGTCGTGAAGAATCTGTGGAGGAGTTTGTA 624

 210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
     |||||||||||||||   ||||||||||||||||||||||||||||||||
 625 CGGCGTAACCTCGGTGATGAGGTTTTTGAGCGCCTGATTGAACCGTTTTG 674

 226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
     |||||||||||||:::||||||||||||||||||||||||||||||||||
 675 TTCAGGTGTTTATGCTGGTGATCCTTCAAAACTGAGCATGAAAGCAGCGT 724

 243 heGlyLysValTrpLysLeuGluGluThrGlyGlySerIleIleGlyGly 259
     |||||||||||||||||||||||::::::|||||||||||||||||||||
 725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGGTGGAAGCATAATAGGTGGT 774

 260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
     |||||||||||||||:::||||||:::::::::|||||||||   |||||
 775 ACTTTTAAGGCAATTCAGGAGAGGAAAAACGCTCCCAAGGCAGAACGAGA 824

 276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
     |||||||||||||||||||:::||||||||||||||||||||||||||||
 825 CCCGCGCCTGCCAAAACCACAGGGCCAAACAGTTGGTTCTTTCAGGAAGG 874

 293 lyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
     |||||||||||||||||:::|||||||||||||||||||||||||||:::
 875 GACTTCGAATGTTGCCAGAAGCAATATCTGCAAGATTAGGTAGCAAAGTT 924

 310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGlyTy 326
     |||||||||||||||||||||:::|||||||||   |||:::||||||||
 925 AAGTTGTCTTGGAAGCTCTCAGGTATCACTAAGCTGGAGAGCGGAGGATA 974

 326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArgS 343
     |:::||||||||||||||||||:::|||:::||||||:::||||||:::|
 975 CAACTTAACATATGAGACTCCAGATGGTTTAGTTTCCGTGCAGAGCAAAA 1024

 343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
     ||:::||||||||||||||||||:::|||||||||::::::|||||||||
1025 GTGTTGTAATGACGGTGCCATCTCATGTTGCAAGTGGTCTCTTGCGCCCT 1074

 360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
     ||||||   :::||||||:::|||||||||::::::||||||||||||||
1075 CTTTCTGAATCTGCTGCAAATGCACTCTCAAAACTATATTACCCACCAGT 1124

 376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
     |:::|||||||:::|||||||||||:::||||||||||||   |||   |
1125 TGCAGCAGTATCTATCTCGTACCCGAAAGAAGCAATCCGAACAGAATGTT 1174

 393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
     ||:::|||||||||||||||||||||||||||||||||||||||||||||
1175 TGATAGATGGTGAACTAAAGGGTTTTGGGCAATTGCATCCACGCACGCAA 1224

 410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1225 GGAGTTGAAACATTAGGAACTATCTACAGCTCCTCACTCTTTCCAAATCG 1274

 426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysA 443
     |||||||   ||||||:::||||||||||||||||||||||||:::   |
1275 CGCACCGCCCGGAAGAATTTTGCTGTTGAACTACATTGGCGGGTCTACAA 1324

 443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValValAsp 459
     ||      ||||||||||||:::|||::::::|||||||||:::||||||
1325 ACACCGGAATTCTGTCCAAGTCTGAAGGTGAGTTAGTGGAAGCAGTTGAC 1374

 460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
     ||||||||||||||||||||||||||||||::::::   |||||||||
1375 AGAGATTTGAGGAAAATGCTAATTAAGCCTAATTCGACCGATCCACTTAA 1424

 476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
      :::||||||||||||||||||||||||||||||||||||||||||||||
1425 ATTAGGAGTTAGGGTATGGCCTCAAGCCATTCCTCAGTTTCTAGTTGGTC 1474

 493 isLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
     ||:::|||   |||:::|||||||||::::::::::::::::::|||
1475 ACTTTGATATCCTTGACACGGCTAAATCATCTCTAACGTCTTCGGGCTAC 1524

 510 GluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
     ||||||||||||||||||||||||||||||:::|||||||||||||||||
1525 GAAGGGCTATTTTTGGGTGGCAATTACGTCGCTGGTGTAGCCTTAGGCCG 1574

 526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
     ||||||||||||||||||||||:::|||   ||||||::::::|||:::|
1575 GTGTGTAGAAGGCGCATATGAAACCGCGATTGAGGTCAACAACTTCATGT 1624

 543 erArgTyrAlaTyrLys 548
     |||||||||||||||||
1625 CACGGTACGCTTACAAG 1641

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-11

seq_documentation_block:
; Sequence 11, Application US/08808931
; Patent No. 5939602
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5939602artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,931
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1847 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
```

```
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  soybean
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-12 (NRRL B-21516)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  55..1683
;      OTHER INFORMATION:  /product= "soybean protox-1"
US-08-808-931-11

alignment_scores:
           Quality: 2062.50            Length:     550
             Ratio:  4.218               Gaps:       7
 Percent Similarity:  88.909   Percent Identity: 73.273

alignment_block:
 US-09-508-418-2 x US-08-808-931-11   ..

 Align seg 1/1  to: US-08-808-931-11  from: 1  to: 1847

   3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
     ::::::|||:::         ||||||:::|||   |||   |||   ||
 102 TCTTCGCCCCTCCCTCCATTCCCCAACCTCTTTCTTCACCTCTCCCACTC 151

  19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerI 36
     |:::                              ||||||
 152 GAAA..............................ATTCCC.........T 162

  36 le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
        |||:::   |||   :::                     ||||||||
 163 CGCTCTCGCCCTAACCCTATT..................CTACGCTGCTC 194

  52 rValAlaLysAspTyrThrVal...ProSerSerAlaValAspGlyGlyP 68
     |:::|||::::::   |||:::   |||   ::::::   |||::::::|
 195 CATTGCGGAGGAATCCACCGCGTCTCCGCCCAAAACCAGAGACTCCGCCC 244

  68 roAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
     ||         :::|||||||||:::||||||:::|||:::|||||||||
 245 CC.........GTGGACTGCGTCGTCGTCGGCGGAGGCGTCAGCGGCCTC 285

  85 CysIleAlaGlnValMetSerAlaAsnTyr......ProAsnLeuMetVa 99
     ||||||||||||::::::::::::::::::         |||::::::||
 286 TGCATCGCCCAGGCCCTCGCCACCAAACACGCCAATGCCAACGTCGTCGT 335

  99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgA 116
     |||||||||||||||||||:::||||||||||||||||||:::|||||||
 336 CACGGAGGCCCGAGACCGCGTCGGCGGCAACATCACCACGATGGAGAGGG 385

 116 spGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 386 ACGGATACCTCTGGGAAGAAGGCCCCAACAGCTTCCAGCCTTCTGATCCA 435

 133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGl 149
     ||||||||||||:::||||||   ||||||||||||:::|||||||||||
 436 ATGCTCACCATGGTGGTGGACAGTGGTTTAAAGGATGAGCTTGTTTTGGG 485

 149 yAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
     |||||||:::|||||||||||||||||||||:::   |||||||||||||
 486 GGATCCTGATGCACCTCGGTTTGTGTTGTGGAACAGGAAGTTGAGGCCGG 535

 166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
     |||||:::|||||||||||||||||||||||||||||||||||||||
 536 TGCCCGGGAAGCTGACTGATTTGCCTTTCTTTGACTTGATGAGCATTGGT 585

 183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
     ||||||:::|||||||||||||||   :::|||:::||||||   |||||
 586 GGCAAAATCAGGGCTGGCTTTGGTGCGCTTGGAATTCGGCCTCCTCCTCC 635

 199 oGlyHisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyG 216
     ||||||||||||||||||||||:::|||||||||||||||||||||   |
 636 AGGTCATGAGGAATCGGTTGAAGAGTTTGTTCGTCGGAACCTTGGTGATG 685

 216 luValPheGluArgLeuIleGluProPheCysSerGlyValTyrValGly 232
     ||||||||||||||||||||||||||||||||||||||||||||:::|||
 686 AGGTTTTTGAACGGTTGATAGAGCCTTTTTGTTCAGGGGTCTATGCAGGC 735

 233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLe 249
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 736 GATCCTTCAAAATTAAGTATGAAAGCAGCATTCGGGAAAGTTTGGAAGCT 785

 249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
     ||||::::::||||||||||||||||||||||||||||||||||||:::|
 786 GGAAAAAAATGGTGGTAGCATTATTGGTGGAACTTTCAAAGCAATACAAG 835

 266 luArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
     |||||:::::::::   |||   |||||||||||||||||||||||||||
 836 AGAGAAATGGAGCTTCAAAACCACCTCGAGATCCGCGTCTGCCAAAACCA 885

 283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
     ||||||||||||||||||||||||||||||||||||   |||||||||||
 886 AAAGGTCAGACTGTTGGATCTTTCCGGAAGGGACTTACCATGTTGCCTGA 935

 299 pAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
     ||||||||||||||||||||||:::|||:::|||||||||||||||||||
 936 TGCAATTTCTGCCAGACTAGGCAACAAAGTAAAGTTATCTTGGAAGCTTT 985

 316 erSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThr 332
     ||||||||:::|||   ::::::|||   |||   |||||||||||||||
 986 CAAGTATTAGTAAACTGGATAGTGGAGAGTACAGTTTGACATATGAAACA 1035

 333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetThrValPr 349
     ||||||||||||||||||||||||   :::::::::|||:::|||:::||
1036 CCAGAAGGAGTGGTTTCTTTGCAGTGCAAAACTGTTGTCCTGACCATTCC 1085

 349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaA 366
     ||||||||||||||||::::::|||||||||||||||:::||||||||||
1086 TTCCTATGTTGCTAGTACATTGCTGCGTCCTCTGTCTGCTGCTGCTGCAG 1135

 366 spAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSer 382
     |||||||||||:::||||||||||||||||||:::||||||:::||||||
1136 ATGCACTTTCAAAGTTTATTACCCTCCAGTTGCTGCAGTTTCCATATCC 1185

 383 TyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLy 399
     ||||||:::||||||||||||:::|||   |||:::||||||||||||||
1186 TATCCAAAAGAAGCTATTAGATCAGAATGCTTGATAGATGGTGAGTTGAA 1235

 399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyT 416
     |||||||||||||||||||||||||:::||||||||||||||||||||||
1236 GGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAACATTAGGAA 1285

 416 hrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
     ||||||||||||||||||||||||||||||||||||||   |||||||||
1286 CTATATACAGCTCATCACTATTCCCCAACCGAGCACCACCTGGAAGGGTT 1335

 433 LeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLy 449
     |||||||||||||||||||||||||||   |||      |||||||||||
1336 CTACTCTTGAATTACATTGGAGGAGCAACTAATACTGGAATTTTATCGAA 1385

 449 sThrGluSerGlnLeuValGluValValAspArgAspLeuArgLysMetL 466
     ||||:::|||:::|||||||||:::|||||||||||||||||||||:::|
1386 GACGGACAGTGAACTTGTGGAAACAGTTGATCGAGATTTGAGGAAAATCC 1435

 466 euIleLysProLysAlaGlnAspProLeuValValGlyValArgValTrp 482
     |||||:::|||:::||||||||||||:::|||||||||||||||:::|||
1436 TTATAAACCCAAATGCCCAGGATCCATTTGTAGTGGGGGTGAGACTGTGG 1485

 483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
     |||||||||||||||||||||||||||||||||||||||   |||:::::
1486 CCTCAAGCTATTCCACAGTTCTTAGTTGGCCATCTTGATCTTCTAGATGT 1535

 499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
     :|||||||||:::::::::::::::|||:::|||||||||||||||||||
1536 TGCTAAAGCTTCTATCAGAAATACTGGGTTTGAAGGGCTCTTCCTTGGGG 1585
```

```
     516 lyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1586 GTAATTATGTGTCTGGTGTTGCCTTGGGACGATGCGTTGAGGGAGCCTAT 1635

     533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
         |||||||||:::||||||:::   ||||||::::::   :::||||||
    1636 GAGGTAGCAGCTGAAGTAAACGATTTTCTCACAAATAGAGTGTACAAA 1683

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-808-323-11

seq_documentation_block:
; Sequence 11, Application US/08808323
; Patent No. 6018105
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Ward, Eric
;    TITLE OF INVENTION:  Promoters from Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase Genes
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6018105artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,323
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1846
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1847 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  soybean
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-12 (NRRL B-21516)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  55..1683
;      OTHER INFORMATION:  /product= "soybean protox-1"
US-08-808-323-11


alignment_scores:
           Quality: 2062.50            Length:    550
```

```
                    Ratio:  4.218                Gaps:      7
       Percent Similarity:  88.909   Percent Identity: 73.273

alignment_block:
  US-09-508-418-2 x US-08-808-323-11   ..

  Align seg 1/1  to: US-08-808-323-11  from: 1  to: 1847

       3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
         ::::::|||:::         ||||||:::|||   |||   |||   ||
     102 TCTTCGCCCCTCCCTCCATTCCCCAACCTCTTTCTTCACCTCTCCCACTC 151

      19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerI 36
         |:::                              ||||||
     152 GAAA..............................ATTCCC.........T 162

      36 le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
            |||:::   |||   :::                     ||||||||
     163 CGCTCTCGCCCTAACCCTATT..................CTACGCTGCTC 194

      52 rValAlaLysAspTyrThrVal...ProSerSerAlaValAspGlyGlyP 68
         |:::|||::::::   |||:::   |||   ::::::   |||::::::|
     195 CATTGCGGAGGAATCCACCGCGTCTCCGCCCAAAACCAGAGACTCCGCCC 244

      68 roAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
         ||         :::|||||||||:::||||||:::|||:::|||||||||
     245 CC.........GTGGACTGCGTCGTCGTCGGCGGAGGCGTCAGCGGCCTC 285

      85 CysIleAlaGlnValMetSerAlaAsnTyr......ProAsnLeuMetVa 99
         ||||||||||||::::::::::::::::::         |||::::::||
     286 TGCATCGCCCAGGCCCTCGCCACCAAACACGCCAATGCCAACGTCGTCGT 335

      99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgA 116
         |||||||||||||||||||:::||||||||||||||||||:::|||||||
     336 CACGGAGGCCCGAGACCGCGTCGGCGGCAACATCACCACGATGGAGAGGG 385

     116 spGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     386 ACGGATACCTCTGGGAAGAAGGCCCCAACAGCTTCCAGCCTTCTGATCCA 435

     133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGl 149
         ||||||||||||:::||||||   ||||||||||||:::|||||||||||
     436 ATGCTCACCATGGTGGTGGACAGTGGTTTAAAGGATGAGCTTGTTTTGGG 485

     149 yAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
         |||||||:::|||||||||||||||||||||:::   |||||||||||||
     486 GGATCCTGATGCACCTCGGTTTGTGTTGTGGAACAGGAAGTTGAGGCCGG 535

     166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
         |||||:::|||||||||||||||||||||||||||||||||||||||
     536 TGCCCGGGAAGCTGACTGATTTGCCTTTCTTTGACTTGATGAGCATTGGT 585

     183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
         ||||||:::|||||||||||||||   :::|||:::||||||   |||||
     586 GGCAAAATCAGGGCTGGCTTTGGTGCGCTTGGAATTCGGCCTCCTCCTCC 635

     199 oGlyHisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyG 216
         ||||||||||||||||||||||:::|||||||||||||||||||||   |
     636 AGGTCATGAGGAATCGGTTGAAGAGTTTGTTCGTCGGAACCTTGGTGATG 685

     216 luValPheGluArgLeuIleGluProPheCysSerGlyValTyrValGly 232
         ||||||||||||||||||||||||||||||||||||||||||||:::|||
     686 AGGTTTTTGAACGGTTGATAGAGCCTTTTTGTTCAGGGGTCTATGCAGGC 735

     233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLe 249
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     736 GATCCTTCAAAATTAAGTATGAAAGCAGCATTCGGGAAAGTTTGGAAGCT 785

     249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
         ||||::::::||||||||||||||||||||||||||||||||||||:::|
     786 GGAAAAAAATGGTGGTAGCATTATTGGTGGAACTTTCAAAGCAATACAAG 835

     266 luArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
```

```
     |||||:::::::::   |||   |||||||||||||||||||||||||||
 836 AGAGAAATGGAGCTTCAAAACCACCTCGAGATCCGCGTCTGCCAAAACCA 885

 283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
     ||||||||||||||||||||||||||||||||||||   |||||||||||
 886 AAAGGTCAGACTGTTGGATCTTTCCGGAAGGGACTTACCATGTTGCCTGA 935

 299 pAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
     ||||||||||||||||||||||:::|||:::|||||||||||||||||||
 936 TGCAATTTCTGCCAGACTAGGCAACAAAGTAAAGTTATCTTGGAAGCTTT 985

 316 erSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThr 332
     ||||||||:::|||   ::::::|||   |||   |||||||||||||||
 986 CAAGTATTAGTAAACTGGATAGTGGAGAGTACAGTTTGACATATGAAACA 1035

 333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetThrValPr 349
     ||||||||||||||||||||||||   :::::::::|||:::|||:::||
1036 CCAGAAGGAGTGGTTTCTTTGCAGTGCAAAACTGTTGTCCTGACCATTCC 1085

 349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaA 366
     ||||||||||||||||::::::|||||||||||||||:::||||||||||
1086 TTCCTATGTTGCTAGTACATTGCTGCGTCCTCTGTCTGCTGCTGCTGCAG 1135

 366 spAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSer 382
     |||||||||||:::||||||||||||||||||:::||||||:::||||||
1136 ATGCACTTTCAAAGTTTATTACCCTCCAGTTGCTGCAGTTTCCATATCC 1185

 383 TyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLy 399
     ||||||:::||||||||||||:::|||   |||:::||||||||||||||
1186 TATCCAAAAGAAGCTATTAGATCAGAATGCTTGATAGATGGTGAGTTGAA 1235

 399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyT 416
     |||||||||||||||||||||||||:::||||||||||||||||||||||
1236 GGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAACATTAGGAA 1285

 416 hrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
     ||||||||||||||||||||||||||||||||||||||   |||||||||
1286 CTATATACAGCTCATCACTATTCCCCAACCGAGCACCACCTGGAAGGGTT 1335

 433 LeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLy 449
     |||||||||||||||||||||||||||   |||      |||||||||||
1336 CTACTCTTGAATTACATTGGAGGAGCAACTAATACTGGAATTTTATCGAA 1385

 449 sThrGluSerGlnLeuValGluValValAspArgAspLeuArgLysMetL 466
     ||||:::|||:::|||||||||:::|||||||||||||||||||||:::|
1386 GACGGACAGTGAACTTGTGGAAACAGTTGATCGAGATTTGAGGAAAATCC 1435

 466 euIleLysProLysAlaGlnAspProLeuValValGlyValArgValTrp 482
     |||||:::|||:::||||||||||||:::|||||||||||||||:::|||
1436 TTATAAACCCAAATGCCCAGGATCCATTTGTAGTGGGGGTGAGACTGTGG 1485

 483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
     |||||||||||||||||||||||||||||||||||||||   |||:::::
1486 CCTCAAGCTATTCCACAGTTCTTAGTTGGCCATCTTGATCTTCTAGATGT 1535

 499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
     :|||||||||:::::::::::::::|||:::|||||||||||||||||||
1536 TGCTAAAGCTTCTATCAGAAATACTGGGTTTGAAGGGCTCTTCCTTGGGG 1585

 516 lyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1586 GTAATTATGTGTCTGGTGTTGCCTTGGGACGATGCGTTGAGGGAGCCTAT 1635

 533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
     |||||||||:::||||||:::   ||||||::::::   :::||||||
1636 GAGGTAGCAGCTGAAGTAAACGATTTTCTCACAAATAGAGTGTACAAA 1683

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-050-603A-11

seq_documentation_block:
; Sequence 11, Application US/09050603A
; Patent No. 6023012
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6023012artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/050,603A
;      FILING DATE:  30-MAR-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1847 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  soybean
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-12 (NRRL B-21516)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  55..1683
;      OTHER INFORMATION:  /product= "soybean protox-1"
US-09-050-603A-11


alignment_scores:
            Quality: 2062.50            Length:    550
              Ratio:   4.218              Gaps:      7
 Percent Similarity:  88.909   Percent Identity: 73.273

alignment_block:
 US-09-508-418-2 x US-09-050-603A-11   ..

 Align seg 1/1  to: US-09-050-603A-11  from: 1  to: 1847

       3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
```

```
     ::::::|||:::         ||||||:::|||   |||   |||   ||
 102 TCTTCGCCCCTCCCTCCATTCCCCAACCTCTTTCTTCACCTCTCCCACTC 151

  19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerI 36
     |:::                              ||||||
 152 GAAA..............................ATTCCC.........T 162

  36 le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
        |||:::   |||   :::                     ||||||||
 163 CGCTCTCGCCCTAACCCTATT..................CTACGCTGCTC 194

  52 rValAlaLysAspTyrThrVal...ProSerSerAlaValAspGlyGlyP 68
     |:::|||::::::   |||:::   |||   ::::::   |||::::::|
 195 CATTGCGGAGGAATCCACCGCGTCTCCGCCCAAAACCAGAGACTCCGCCC 244

  68 roAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
     ||         :::|||||||||:::||||||:::|||:::|||||||||
 245 CC.........GTGGACTGCGTCGTCGTCGGCGGAGGCGTCAGCGGCCTC 285

  85 CysIleAlaGlnValMetSerAlaAsnTyr......ProAsnLeuMetVa 99
     ||||||||||||::::::::::::::::::         |||::::::||
 286 TGCATCGCCCAGGCCCTCGCCACCAAACACGCCAATGCCAACGTCGTCGT 335

  99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgA 116
     |||||||||||||||||||:::||||||||||||||||||:::|||||||
 336 CACGGAGGCCCGAGACCGCGTCGGCGGCAACATCACCACGATGGAGAGGG 385

 116 spGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 386 ACGGATACCTCTGGGAAGAAGGCCCCAACAGCTTCCAGCCTTCTGATCCA 435

 133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGl 149
     ||||||||||||:::||||||   ||||||||||||:::|||||||||||
 436 ATGCTCACCATGGTGGTGGACAGTGGTTTAAAGGATGAGCTTGTTTTGGG 485

 149 yAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
     |||||||:::|||||||||||||||||||||:::   |||||||||||||
 486 GGATCCTGATGCACCTCGGTTTGTGTTGTGGAACAGGAAGTTGAGGCCGG 535

 166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
     |||||:::|||||||||||||||||||||||||||||||||||||||
 536 TGCCCGGGAAGCTGACTGATTTGCCTTTCTTTGACTTGATGAGCATTGGT 585

 183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
     ||||||:::|||||||||||||||   :::|||:::||||||   |||||
 586 GGCAAAATCAGGGCTGGCTTTGGTGCGCTTGGAATTCGGCCTCCTCCTCC 635

 199 oGlyHisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyG 216
     ||||||||||||||||||||||:::|||||||||||||||||||||   |
 636 AGGTCATGAGGAATCGGTTGAAGAGTTTGTTCGTCGGAACCTTGGTGATG 685

 216 luValPheGluArgLeuIleGluProPheCysSerGlyValTyrValGly 232
     ||||||||||||||||||||||||||||||||||||||||||||:::|||
 686 AGGTTTTTGAACGGTTGATAGAGCCTTTTTGTTCAGGGGTCTATGCAGGC 735

 233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLe 249
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 736 GATCCTTCAAAATTAAGTATGAAAGCAGCATTCGGGAAAGTTTGGAAGCT 785

 249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
     ||||::::::||||||||||||||||||||||||||||||||||||:::|
 786 GGAAAAAAATGGTGGTAGCATTATTGGTGGAACTTTCAAAGCAATACAAG 835

 266 luArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
     |||||:::::::::   |||   |||||||||||||||||||||||||||
 836 AGAGAAATGGAGCTTCAAAACCACCTCGAGATCCGCGTCTGCCAAAACCA 885

 283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
     ||||||||||||||||||||||||||||||||||||   |||||||||||
 886 AAAGGTCAGACTGTTGGATCTTTCCGGAAGGGACTTACCATGTTGCCTGA 935

 299 pAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
     ||||||||||||||||||||||:::|||:::|||||||||||||||||||
```

```
 936 TGCAATTTCTGCCAGACTAGGCAACAAAGTAAAGTTATCTTGGAAGCTTT 985

 316 erSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThr 332
     ||||||||:::|||   ::::::|||   |||   |||||||||||||||
 986 CAAGTATTAGTAAACTGGATAGTGGAGAGTACAGTTTGACATATGAAACA 1035

 333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetThrValPr 349
     ||||||||||||||||||||||||   :::::::::|||:::|||:::||
1036 CCAGAAGGAGTGGTTTCTTTGCAGTGCAAAACTGTTGTCCTGACCATTCC 1085

 349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaA 366
     ||||||||||||||||::::::|||||||||||||||:::||||||||||
1086 TTCCTATGTTGCTAGTACATTGCTGCGTCCTCTGTCTGCTGCTGCTGCAG 1135

 366 spAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSer 382
     |||||||||||:::||||||||||||||||||:::||||||:::||||||
1136 ATGCACTTTCAAAGTTTTATTACCCTCCAGTTGCTGCAGTTTCCATATCC 1185

 383 TyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLy 399
     ||||||:::||||||||||||:::|||   |||:::||||||||||||||
1186 TATCCAAAAGAAGCTATTAGATCAGAATGCTTGATAGATGGTGAGTTGAA 1235

 399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyT 416
     |||||||||||||||||||||||||:::||||||||||||||||||||||
1236 GGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAACATTAGGAA 1285

 416 hrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
     ||||||||||||||||||||||||||||||||||||||   |||||||||
1286 CTATATACAGCTCATCACTATTCCCCAACCGAGCACCACCTGGAAGGGTT 1335

 433 LeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLy 449
     |||||||||||||||||||||||||||   |||      |||||||||||
1336 CTACTCTTGAATTACATTGGAGGAGCAACTAATACTGGAATTTTATCGAA 1385

 449 sThrGluSerGlnLeuValGluValValAspArgAspLeuArgLysMetL 466
     ||||:::|||:::|||||||||:::|||||||||||||||||||||:::|
1386 GACGGACAGTGAACTTGTGGAAACAGTTGATCGAGATTTGAGGAAAATCC 1435

 466 euIleLysProLysAlaGlnAspProLeuValValGlyValArgValTrp 482
     |||||:::|||:::||||||||||||:::|||||||||||||||:::|||
1436 TTATAAACCCAAATGCCCAGGATCCATTTGTAGTGGGGGTGAGACTGTGG 1485

 483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
     |||||||||||||||||||||||||||||||||||||||   |||:::::
1486 CCTCAAGCTATTCCACAGTTCTTAGTTGGCCATCTTGATCTTCTAGATGT 1535

 499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
     :|||||||||:::::::::::::::|||:::|||||||||||||||||||
1536 TGCTAAAGCTTCTATCAGAAATACTGGGTTTGAAGGGCTCTTCCTTGGGG 1585

 516 lyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1586 GTAATTATGTGTCTGGTGTTGCCTTGGGACGATGCGTTGAGGGAGCCTAT 1635

 533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
     |||||||||:::||||||:::   ||||||::::::   :::||||||
1636 GAGGTAGCAGCTGAAGTAAACGATTTTCTCACAAATAGAGTGTACAAA 1683

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-11

seq_documentation_block:
; Sequence 11, Application US/09102420B
; Patent No. 6084155
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
;    TITLE OF INVENTION:  OXIDASE ("PROTOX")
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6084155artis Corporation
```

```
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/102,420B
;      FILING DATE:  22-JUN-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/059,164
;      FILING DATE:  13-APR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/050,603
;      FILING DATE:  30-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/126,430
;      FILING DATE:  11-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/808,931
;      FILING DATE:  28-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/472,028
;      FILING DATE:  06-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847/CIP3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1847 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  soybean
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-12 (NRRL B-21516)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  55..1683
;      OTHER INFORMATION:  /product= "soybean protox-1"
US-09-102-420B-11


alignment_scores:
           Quality: 2062.50            Length:     550
             Ratio:  4.218               Gaps:       7
 Percent Similarity:  88.909   Percent Identity: 73.273

alignment_block:
 US-09-508-418-2 x US-09-102-420B-11   ..

 Align seg 1/1  to: US-09-102-420B-11  from: 1  to: 1847
```

```
  3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
    ::::::|||:::         ||||||:::|||   |||   |||   ||
102 TCTTCGCCCCTCCCTCCATTCCCCAACCTCTTTCTTCACCTCTCCCACTC 151

 19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerI 36
    |:::                              ||||||
152 GAAA..............................ATTCCC.........T 162

 36 le.SerLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
       |||:::   |||   :::                     ||||||||
163 CGCTCTCGCCCTAACCCTATT..................CTACGCTGCTC 194

 52 rValAlaLysAspTyrThrVal...ProSerSerAlaValAspGlyGlyP 68
    |:::|||::::::   |||:::   |||   ::::::   |||::::::|
195 CATTGCGGAGGAATCCACCGCGTCTCCGCCCAAAACCAGAGACTCCGCCC 244

 68 roAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
    ||         :::|||||||||:::||||||:::|||:::|||||||||
245 CC.........GTGGACTGCGTCGTCGTCGGCGGAGGCGTCAGCGGCCTC 285

 85 CysIleAlaGlnValMetSerAlaAsnTyr......ProAsnLeuMetVa 99
    ||||||||||||::::::::::::::::::         |||::::::||
286 TGCATCGCCCAGGCCCTCGCCACCAAACACGCCAATGCCAACGTCGTCGT 335

 99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgA 116
    |||||||||||||||||||:::||||||||||||||||||:::|||||||
336 CACGGAGGCCCGAGACCGCGTCGGCGGCAACATCACCACGATGGAGAGGG 385

116 spGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
    ||||||||||||||||||||||||||||||||||||||||||||||||||
386 ACGGATACCTCTGGGAAGAAGGCCCCAACAGCTTCCAGCCTTCTGATCCA 435

133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGl 149
    ||||||||||||:::||||||   ||||||||||||:::|||||||||||
436 ATGCTCACCATGGTGGTGGACAGTGGTTTAAAGGATGAGCTTGTTTTGGG 485

149 yAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
    |||||||:::|||||||||||||||||||||:::   |||||||||||||
486 GGATCCTGATGCACCTCGGTTTGTGTTGTGGAACAGGAAGTTGAGGCCGG 535

166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
    |||||:::|||||||||||||||||||||||||||||||||||||||
536 TGCCCGGGAAGCTGACTGATTTGCCTTTCTTTGACTTGATGAGCATTGGT 585

183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProPr 199
    ||||||:::|||||||||||||||   :::|||:::||||||   |||||
586 GGCAAAATCAGGGCTGGCTTTGGTGCGCTTGGAATTCGGCCTCCTCCTCC 635

199 oGlyHisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyG 216
    ||||||||||||||||||||||:::|||||||||||||||||||||   |
636 AGGTCATGAGGAATCGGTTGAAGAGTTTGTTCGTCGGAACCTTGGTGATG 685

216 luValPheGluArgLeuIleGluProPheCysSerGlyValTyrValGly 232
    ||||||||||||||||||||||||||||||||||||||||||||:::|||
686 AGGTTTTTGAACGGTTGATAGAGCCTTTTTGTTCAGGGGTCTATGCAGGC 735

233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLe 249
    ||||||||||||||||||||||||||||||||||||||||||||||||||
736 GATCCTTCAAAATTAAGTATGAAAGCAGCATTCGGGAAAGTTTGGAAGCT 785

249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
    ||||::::::||||||||||||||||||||||||||||||||||||:::|
786 GGAAAAAAATGGTGGTAGCATTATTGGTGGAACTTTCAAAGCAATACAAG 835

266 luArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
    |||||:::::::::   |||   |||||||||||||||||||||||||||
836 AGAGAAATGGAGCTTCAAAACCACCTCGAGATCCGCGTCTGCCAAAACCA 885

283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
    ||||||||||||||||||||||||||||||||||||   |||||||||||
886 AAAGGTCAGACTGTTGGATCTTTCCGGAAGGGACTTACCATGTTGCCTGA 935
```

```
 299 pAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuS 316
     ||||||||||||||||||||||:::|||:::|||||||||||||||||||
 936 TGCAATTTCTGCCAGACTAGGCAACAAAGTAAAGTTATCTTGGAAGCTTT 985

 316 erSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThr 332
     ||||||||:::|||   ::::::|||   |||   |||||||||||||||
 986 CAAGTATTAGTAAACTGGATAGTGGAGAGTACAGTTTGACATATGAAACA 1035

 333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetThrValPr 349
     ||||||||||||||||||||||||   :::::::::|||:::|||:::||
1036 CCAGAAGGAGTGGTTTCTTTGCAGTGCAAAACTGTTGTCCTGACCATTCC 1085

 349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaA 366
     ||||||||||||||||::::::|||||||||||||||:::||||||||||
1086 TTCCTATGTTGCTAGTACATTGCTGCGTCCTCTGTCTGCTGCTGCTGCAG 1135

 366 spAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSer 382
     |||||||||||:::||||||||||||||||||:::||||||:::||||||
1136 ATGCACTTTCAAAGTTTATTACCCTCCAGTTGCTGCAGTTTCCATATCC 1185

 383 TyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLy 399
     ||||||:::||||||||||||:::|||   |||:::||||||||||||||
1186 TATCCAAAAGAAGCTATTAGATCAGAATGCTTGATAGATGGTGAGTTGAA 1235

 399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyT 416
     |||||||||||||||||||||||||:::||||||||||||||||||||||
1236 GGGGTTTGGTCAATTGCATCCACGTAGCCAAGGAGTGGAAACATTAGGAA 1285

 416 hrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
     ||||||||||||||||||||||||||||||||||||||   |||||||||
1286 CTATATACAGCTCATCACTATTCCCCAACCGAGCACCACCTGGAAGGGTT 1335

 433 LeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLy 449
     |||||||||||||||||||||||||||   |||      |||||||||||
1336 CTACTCTTGAATTACATTGGAGGAGCAACTAATACTGGAATTTTATCGAA 1385

 449 sThrGluSerGlnLeuValGluValValAspArgAspLeuArgLysMetL 466
     ||||:::|||:::||||||||||||:::||||||||||||||||||:::|
1386 GACGGACAGTGAACTTGTGGAAACAGTTGATCGAGATTTGAGGAAAATCC 1435

 466 euIleLysProLysAlaGlnAspProLeuValValGlyValArgValTrp 482
     |||||:::|||:::||||||||||||:::|||||||||||||||:::|||
1436 TTATAAACCCAAATGCCCAGGATCCATTTGTAGTGGGGGTGAGACTGTGG 1485

 483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
     |||||||||||||||||||||||||||||||||||||||   |||:::::
1486 CCTCAAGCTATTCCACAGTTCTTAGTTGGCCATCTTGATCTTCTAGATGT 1535

 499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
     :|||||||||:::::::::::::::|||:::|||||||||||||||||||
1536 TGCTAAAGCTTCTATCAGAAATACTGGGTTTGAAGGGCTCTTCCTTGGGG 1585

 516 lyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1586 GTAATTATGTGTCTGGTGTTGCCTTGGGACGATGCGTTGAGGGAGCCTAT 1635

 533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
     |||||||||:::||||||:::   ||||||::::::   :::||||||
1636 GAGGTAGCAGCTGAAGTAAACGATTTTCTCACAAATAGAGTGTACAAA 1683

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-19

seq_documentation_block:
; Sequence 19, Application US/08808931
; Patent No. 5939602
;  GENERAL INFORMATION:
;    APPLICANT:  Volrath, Sandra
;    APPLICANT:  Johnson, Marie
;    APPLICANT:  Potter, Sharon
;    APPLICANT:  Ward, Eric
;    APPLICANT:  Heifetz, Peter
;    TITLE OF INVENTION:  DNA Molecules Encoding Plant
;    TITLE OF INVENTION:  Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
```

```
;    TITLE OF INVENTION:  Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5939602artis Corporation
;      STREET:  520 White Plains Road, P.O. Box 2005
;      CITY:  Tarrytown
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10591-9005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/808,931
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/012,705
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/013,612
;      FILING DATE:  28-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/020,003
;      FILING DATE:  21-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  CGC 1847
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (919) 541-8587
;      TELEFAX:  (919) 541-8689
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1784 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Brassica napus (rape)
;    IMMEDIATE SOURCE:
;      CLONE:  pWDC-17 (NRRL B-21615)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  47..1654
;      OTHER INFORMATION:  /product= "Rape Protox-1 coding
;      OTHER INFORMATION:  region"
US-08-808-931-19


alignment_scores:
           Quality:  2006.50              Length:    535
             Ratio:  4.189                  Gaps:      4
 Percent Similarity:  89.533   Percent Identity:  73.084

alignment_block:
 US-09-508-418-2 x US-08-808-931-19   ..

 Align seg 1/1  to: US-08-808-931-19  from: 1  to: 1784

      21 ProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerIleSe 37
         |||      |||||||:::               |||||||||:::
      68 CCGCAGCCATTCCTATCG...............CCATTCTCAAATCCATT 102

      37 rLysArgAsnSerValAsnCysAsnGlyTrpArgThrArgCysSerValA 54
             |||   |||         :::      :::   ||||||||||||:
     103 TCCTCGG...TCGCGTCCCTACAAGCCTCTCAACCTCCGTTGCTCCGTAT 149
```

```
  54 laLysAspTyrThrValProSerSerAlaValAspGlyGlyProAlaAla 70
     ::         :::|||   ||||||:::::::::||||||   ::::::
 150 CCGGTGGATCCGTCGTCGGCTCTTCTACAATCGAAGGCGGAGGAGGAGGT 199

  71 Glu.........LeuAspCysValIleValGlyAlaGlyIleSerGlyLe 84
     :::            ||||||||||||||||||:::||||||||||||||
 200 AAAACCGTCACGGCGGACTGCGTGATCGTCGGCGGAGGAATCAGCGGCCT 249

  84 uCysIleAlaGlnValMetSerAlaAsnTyrPro............AsnL 97
     |||||||||||||::::::   :::::::::|||            |||:
 250 GTGCATTGCGCAAGCGCTCGTGACGAAGCACCCAGACGCTGCAAAGAATG 299

  97 euMetValThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrVal 113
     ::|||||||||||||||:::||||||:::||||||||||||   |||
 300 TGATGGTGACGGAGGCGAAGGACCGTGTGGAGGGAATATCATCACGCGA 349

 114 GluArgAspGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnProSe 130
     |||::::::|||:::|||||||||||||||||||||||||||||||||||
 350 GAGGAGCAAGGGTTTCTATGGGAAGAAGGTCCCAATAGCTTTCAGCCGTC 399

 130 rAspProMetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuV 147
     |||||||||||||||||||:::||||||   |||||||||||||||||||
 400 TGATCCTATGCTCACTATGGTGGTAGATAGTGGTTTGAAAGATGATCTAG 449

 147 alLeuGlyAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeu 163
     ||||||||||||||:::|||||||||||||||||||||:::|||||||||
 450 TCTTGGGAGATCCTACTGCTCCGAGGTTTGTGTTGTGGAATGGGAAGCTG 499

 164 ArgProValProSerLysLeuThrAspLeuProPhePheAspLeuMetSe 180
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 500 AGGCCGGTTCCGTCGAAGCTAACTGACTTGCCTTTCTTTGACTTGATGAG 549

 180 rIleProGlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProS 197
     ||||   ||||||:::|||||||||||||||   ||||||:::|||||||
 550 TATTGGAGGGAAGATTAGAGCTGGGTTTGGTGCCATTGGTATTCGACCTT 599

 197 erProProGlyHisGluGluSerValGluGlnPheValArgArgAsnLeu 213
     |||||||||||:::|||||||||||||||:::||||||||||||||||||
 600 CACCTCCGGGTCGTGAGGAATCAGTGGAAGAGTTTGTAAGGCGTAATCTT 649

 214 GlyGlyGluValPheGluArgLeuIleGluProPheCysSerGlyValTy 230
     |||   ||||||||||||||||||||||||||||||||||||||||||||
 650 GGTGATGAGGTTTTTGAGCGCTTGATTGAACCCTTTTGCTCAGGTGTTTA 699

 230 rValGlyAspProSerLysLeuSerMetLysAlaAlaPheGlyLysValT 247
     |:::|||||||||:::||||||||||||||||||||||||||||||||||
 700 TGCGGGAGATCCTGCGAAACTGAGTATGAAAGCAGCTTTTGGGAAGGTTT 749

 247 rpLysLeuGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAla 263
     ||||||||||||||:::|||||||||||||||||||||:::|||||||||
 750 GGAAGCTAGAGGAGAATGGTGGGAGCATCATTGGTGGTGCTTTTAAGGCA 799

 264 IleLysGluArgSerSerThrProLysAlaProArgAspProArgLeuPr 280
     |||:::   ::::::::::::||||||:::   |||||||||||||||||
 800 ATTCAAGCGAAAAATAAAGCTCCCAAGACAACCCGAGATCCGCGTCTGCC 849

 280 oLysProLysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetL 297
     |||||||||||||||||||||||||||||||||||||||||||   ||||
 850 AAAGCCAAAGGGCCAAACTGTTGGTTCTTTCAGGAAAGGACTCACAATGC 899

 297 euProAspAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrp 313
     |||||:::|||||||||||||||||||||:::|||:::|||:::||||||
 900 TGCCAGAGGCAATCTCCGCAAGGTTGGGTGACAAGGTGAAAGTTTCTTGG 949

 314 LysLeuSerSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTy 330
     |||||||||||||||||||||      :::|||   |||   ||||||||
 950 AAGCTCTCAAGTATCACTAAGCTGGCCAGCGGAGAATATAGCTTAACTTA 999

 330 rGluThrProGluGlyValValSerLeuGlnSerArgSerIleValMetT 347
     ||||||||||||||||:::|||::::::||||||:::|||:::|||||||
1000 CGAAACTCCGGAGGGTATAGTCACTGTACAGAGCAAAAGTGTAGTGATGA 1049

 347 hrValProSerTyrValAlaSerAsnIleLeuArgProLeuSerValAla 363
     |||||||||||:::|||||||||::::::|||||||||||||||   :::
1050 CTGTGCCATCTCATGTTGCTAGTAGTCTCTTGCGCCCTCTCTCTGATTCT 1099

 364 AlaAlaAspAlaLeuSerAsnPheTyrTyrProProValGlyAlaValTh 380
     ||||||:::|||||||||::::::|||||||||||||||:::||||||::
1100 GCAGCTGAAGCGCTCTCAAAACTCTACTATCCGCCAGTTGCAGCCGTATC 1149

 380 rIleSerTyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyG 397
     :|||||||||   :::||||||||||||:::|||   |||:::|||||||
1150 CATCTCATACGCGAAAGAAGCAATCCGAAGCGAATGCTTAATAGATGGTG 1199

 397 luLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThr 413
     ||||||||||||||||||||||||||||||||||||||   |||||||||
1200 AACTAAAAGGGTTCGGCCAGTTGCATCCACGCACGCAAAAAGTGGAAACT 1249

 414 LeuGlyThrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGl 430
     |||||||||||||||||||||||||||||||||||||||||||||   ||
1250 CTTGGAACAATATACAGTTCATCGCTCTTTCCCAACCGAGCACCGCCTGG 1299

 430 yArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleL 447
     ||||||||||||||||||||||||||||||||||   |||      ||||
1300 AAGAGTATTGCTATTGAACTACATCGGTGGAGCTACCAACACTGGGATCT 1349

 447 euSerLysThrGluSerGlnLeuValGluValValAspArgAspLeuArg 463
     ||||||||:::|||::::::|||||||||:::||||||||||||||||||
1350 TATCAAAGTCGGAAGGTGAGTTAGTGGAAGCAGTAGATAGAGACTTGAGG 1399

 464 LysMetLeuIleLysProLysAlaGlnAspProLeuValValGlyValAr 480
     ||||||||||||||||||::::::   ||||||||||||:::||||||::
1400 AAGATGCTGATAAAGCCAAGCTCGACCGATCCACTTGTACTTGGAGTAAA 1449

 480 gValTrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrL 497
     ::::|||||||||||||||||||||||||||:::||||||:::|||   :
1450 ATTATGGCCTCAAGCCATTCCTCAGTTTCTGATAGGTCACATTGATTTGG 1499

 497 euSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPhe 513
     ::::::::|||||||||:::::::::::::::|||   ||||||||||||
1500 TAGACGCAGCGAAAGCATCGCTCTCGTCATCTGGTCATGAGGGCTTATTC 1549

 514 LeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGl 530
     ||||||||||||||||||:::|||||||||||||||||||||||||||||
1550 TTGGGTGGAAATTACGTTGCCGGTGTAGCATTGGGTCGGTGTGTGGAAGG 1599

 530 yAlaTyrGluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaT 547
     ||||||||||:::|||::::::|||:::   |||:::|||||||||||||
1600 TGCTTATGAAACTGCAACCCAAGTGAATGATTTCATGTCAAGGTATGCTT 1649

 547 yrLys 548
     |||||
1650 ACAAG 1654
```